

# Essays on asset market comovements

Citation for published version (APA):

Piplack, J. (2008). *Essays on asset market comovements*. [Doctoral Thesis, Maastricht University]. Maastricht University Press. <https://doi.org/10.26481/dis.20080926jp>

**Document status and date:**

Published: 01/01/2008

**DOI:**

[10.26481/dis.20080926jp](https://doi.org/10.26481/dis.20080926jp)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

ESSAYS ON  
ASSET MARKET COMOVEMENTS

JAN PIPLACK

© Jan Piplack, 2008

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form, or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing from the author.

This book was typeset by the author using L<sup>A</sup>T<sub>E</sub>X.

Published by Universitaire Pers Maastricht

ISBN: XX-XXXX-XXX-X

Printed in the Netherlands by Datawyse Maastricht

# ESSAYS ON ASSET MARKET COMOVEMENTS

PROEFSCHRIFT

ter verkrijging van de graad van doctor  
aan de Universiteit Maastricht,  
op gezag van de Rector Magnificus,  
Prof. mr. G.P.M.F. Mols,  
volgens het besluit van het College van Decanen,  
in het openbaar te verdedigen op  
vrijdag 26 september 2008 om 10.00 uur

door

Jan Piplack

**Promotores:**

Prof. dr. B. Candelon

Prof. dr. J. Muysken

**Beoordelingscommissie:**

Prof. dr. C.C.P. Wolff (voorzitter)

Prof. dr. Y.-W. Cheung (University of California in Santa Cruz)

Prof. dr. J.-P. Urbain

Dit onderzoek werd gefinancierd door de Maastricht Research School of Economics of Technology and Organizations (METEOR).

When the number of factors coming into play in a phenomenological complex is too large, scientific method in most cases fails us. One need only think of the weather, in which case prediction even for a few days ahead is impossible. Nevertheless no one doubts that we are confronted with a causal connection whose causal components are in the main known to us. Occurrences in this domain are beyond the reach of exact prediction because of the variety of factors in operation, not because of any lack of order in nature. (Albert Einstein (1954, pp.70-71))



---

## ACKNOWLEDGEMENTS

Working on my dissertation was fun most of the time. Anyway, getting to the end of the project I started thinking about the last years more extensively. All of a sudden I realized that this stage of the journey is really over and that it probably started way earlier than September 2004. Some time before I finished high school my parents and I talked about options for the future and I remember my father pointing out that there is no Ph.D. in the family yet. I guess this statement had its share in this book and I feel somewhat relieved now. In the last four years I further realized that the job as a researcher/Ph.D. candidate can sometimes be difficult. When you speak to people outside the academic world you have to defend and explain yourself because many consider you still a student, who does unnecessary things. At times they might have a point. Within academia you have to defend your work against criticism, which often is warranted. Occasionally this makes you feel alone. Even though doing research mostly is a lonely job, there have been many people assisting me on the way, which made writing this dissertation an interesting learning process for me. Some of which I would like to thank explicitly here.

I am grateful to my advisor Bertrand Candelon for his encouragement, guidance, and at the same time freedom he gave me to pursue my own ideas. Bertrand always had an open office door and made time when I needed his opinion and advice. It has been a pleasure to work with him and I hope that we manage to stay in professional and private contact. My second advisor Joan Muysken I would like to thank for his support and discussions which we had especially in the last phase of my project. I really appreciate the comments he had on my work.

I owe thanks to Stefan Straetmans and Michel Beine for their cooperation in some of the chapters of this dissertation. Moreover, I thank the members of the reading and assessment committee: Christian Wolff, Yin-Wong Cheung and Jean-Pierre Urbain, for their careful reading and constructive comments. I feel



## ACKNOWLEDGEMENTS

---

especially indebted to Yin-Wong who invited me to spend seven months at the University of California in Santa Cruz in 2007. I benefited very much from his support and comments on my dissertation. Also Jean-Pierre I owe a lot for his constructive criticism and intellectual challenges he gave me during the last four years.

Being able to share ideas and thoughts with people who are in the same position as you are, or short fellow-“sufferers”, was very important to me. First to mention here is my paranymp, first officemate, and by now good friend Katharina Raabe who I shared the room with during the major part of my time in Maastricht. I am still impressed by her self-discipline, energy and working-times. During some months I tried to keep up with her, but I had to acknowledge my limitations. I really miss our “muffin-breaks” and office talks about life and sailing. Further, I would like to point out my special thanks to Jeroen van den Berg (my second officemate), Nils Holinski, Stefan Kühn, Kasper Leufkens (Manni), and Stephan Smeekes who contributed a lot to a pleasant atmosphere at work. Jeroen and Kasper, for example, rescued me during my stay in California when my office computer in Maastricht met it’s maker. They spent days from their time fixing it so that I could use it again. Stephan helped me a lot with his mathematical mind on parts of my dissertation where my mathematical capabilities attained their limits. I will definitely miss our joint daily lunches and discussions about football.

I would like to thank all my other colleagues from Maastricht University where my special thanks go to Sylvia Beenen, Fleur Keune, and Silvana de Sanctis. Without their helpfulness and organizational and administrative support my life as a Ph.D. student would have been much harder.

For helping me a lot with the design and technical work related to the cover of this book, I thank Stephanie Blaschka. She saved me a lot of computer-frustration.

The most I owe to my parents Birgitta and Helmut Piplack. They have altruistically and generously supported me my whole life and also during the last four years where they had to listen to some doubts and scepticism from my side. Their mental encouragement has been vital to me. I also would like to thank Lotte for going through this together with me and for always helping me with my Dutch language problems like writing the summary in this book.

Finally I would like to thank the Maastricht Research School of Economics of Technology and Organizations (METEOR) for financing my research.

---

# CONTENTS

<b>Acknowledgements</b>	<b>vii</b>
<b>Contents</b>	<b>ix</b>
<b>List of Tables</b>	<b>xii</b>
<b>List of Figures</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The general idea of this work . . . . .	1
1.2 Measures and sources of asset market linkages . . . . .	2
1.3 The importance of asset markets . . . . .	5
1.4 Aim of the study . . . . .	7
1.5 Structure of the thesis . . . . .	9
<b>2 Comovements of Returns and Volatility in International Stock Markets: A High-Frequency Approach</b>	<b>11</b>
2.1 Introduction . . . . .	11
2.2 Methodology . . . . .	14
2.2.1 The basic setup . . . . .	14
2.2.2 Realized measures and jump test statistic . . . . .	15
2.2.3 Common factors in the continuous component . . . . .	18
2.2.4 Co-jumps . . . . .	20
2.3 Empirical results . . . . .	21
2.3.1 Continuous and jump component . . . . .	21
2.3.2 Common factors in the volatility . . . . .	27
2.3.3 Co-jumps . . . . .	32

ix

## CONTENTS

---

2.3.4	Close-open returns and jumps . . . . .	35
2.3.5	Close-open returns across countries . . . . .	38
2.4	Conclusions . . . . .	42
<b>3</b>	<b>Multivariate Stock Market Synchronization in Small Samples: The Case of East-Asia</b>	<b>43</b>
3.1	Introduction . . . . .	43
3.2	Strong multivariate synchronization . . . . .	46
3.3	Small sample properties of the SMS( $\rho_0$ ) test . . . . .	49
3.4	Consistency of the block bootstrap method . . . . .	58
3.5	Empirical application . . . . .	62
3.6	Conclusions . . . . .	69
<b>4</b>	<b>On Measuring Synchronization of Bulls and Bears: The Case of East Asia</b>	<b>71</b>
4.1	Introduction . . . . .	71
4.2	Methodology . . . . .	74
4.2.1	Strong multivariate synchronization of degree $\rho_0$ . . . . .	74
4.2.2	Structural breaks in multivariate synchronization . . . . .	76
4.3	Monte Carlo experiments . . . . .	79
4.4	Empirical evidence . . . . .	84
4.5	Conclusions . . . . .	92
4.6	Appendix: Block bootstrap algorithm . . . . .	95
<b>5</b>	<b>Comovements of Different Asset Classes During Market Stress</b>	<b>97</b>
5.1	Introduction . . . . .	97
5.2	Asset linkages: Theory . . . . .	99
5.2.1	Conditional tail probabilities . . . . .	99
5.2.2	Conditionally expected extreme events . . . . .	100
5.3	Estimation of the linkage indicators . . . . .	101
5.4	Hypothesis testing . . . . .	105
5.4.1	Time variation . . . . .	105
5.4.2	Cross-sectional variation . . . . .	108
5.5	Extreme asset linkage results: Stocks, bonds, T-bills, gold . . . . .	108
5.5.1	Univariate results . . . . .	110
5.5.2	Bivariate results . . . . .	116
5.6	Conclusions . . . . .	134
5.7	Appendix . . . . .	136

<b>6</b>	<b>Estimating and Forecasting Asset Volatility and Its Volatility: A Markov-Switching Range Model</b>	<b>137</b>
6.1	Introduction . . . . .	137
6.2	Theoretical foundations . . . . .	140
6.2.1	Theory of the range as a volatility estimator . . . . .	140
6.2.2	Extension to non i.i.d. case . . . . .	141
6.2.3	Monte Carlo experiment . . . . .	142
6.3	Methodology . . . . .	144
6.3.1	The path to the model . . . . .	144
6.3.2	The model . . . . .	148
6.3.3	Estimation . . . . .	150
6.3.4	Smoothed regime inference . . . . .	153
6.4	Application and results . . . . .	155
6.4.1	The data . . . . .	155
6.4.2	Estimation results . . . . .	160
6.4.3	Comovements of volatility across asset markets . . . . .	168
6.4.4	Forecasting performance . . . . .	175
6.5	Conclusions . . . . .	178
	<b>Bibliography</b>	<b>181</b>
	<b>Nederlandse Samenvatting (Summary in Dutch)</b>	<b>197</b>
	<b>Short Curriculum Vitae</b>	<b>205</b>

---

## LIST OF TABLES

2.1	Descriptive statistics . . . . .	27
2.2	Co-jumps probability in % . . . . .	33
2.3	Extreme close-open returns and jumps . . . . .	36
2.4	Co-extremes in close-open returns (unadjusted) . . . . .	39
2.5	Co-extremes in close-open returns (adjusted) . . . . .	40
3.1	Full sample cyclical stock market synchronization in South East Asia: Estimates and homogeneity tests . . . . .	65
3.2	Subsample cyclical stock market synchronization in South East Asia: Estimates and homogeneity tests . . . . .	68
4.1	Small sample size of stability test: Uncorrected and bootstrap-corrected . . . . .	81
4.2	Small sample (size-corrected) power of bootstrap stability test . . . . .	83
4.3	Full sample market synchronization: Estimates and homogeneity tests . . . . .	87
4.4	Testing for structural change in multivariate synchronization index . . . . .	89
4.5	Subsample cyclical stock market synchronization in South East Asia: Estimates and homogeneity tests . . . . .	90
5.1	Descriptive statistics for daily US asset returns . . . . .	110
5.2	Minima, maxima, tail index, and univariate tail estimates for daily US asset returns . . . . .	111
5.3	Univariate results for stability test for daily US asset returns and squared returns . . . . .	115

## LIST OF TABLES

---

5.4	Bivariate results: Full sample . . . . .	121
5.5	Bivariate results: Stability test . . . . .	123
5.6	Bivariate results: Squared returns . . . . .	124
5.7	Bivariate results: Stability test squared returns . . . . .	125
6.1	Descriptive statistics . . . . .	156
6.2	Estimation results for weekly S&P500 data only allowing the constant ( $\mu$ ) to change . . . . .	162
6.3	Results for weekly S&P500 data allowing all mean equation parameters to change . . . . .	166
6.4	Estimation results for different weekly asset markets . . . . .	169
6.5	Point forecast comparison . . . . .	176

---

## LIST OF FIGURES

2.1	UK: Continuous and jump component . . . . .	23
2.2	US: Continuous and jump component . . . . .	23
2.3	The Netherlands: Continuous and jump component . . . . .	24
2.4	Italy: Continuous and jump component . . . . .	24
2.5	Germany: Continuous and jump component . . . . .	25
2.6	France: Continuous and jump component . . . . .	25
2.7	Austria: Continuous and jump component . . . . .	26
2.8	Switzerland: Continuous and jump component . . . . .	26
2.9	Common factors CC: Europe core, US, Austria and Switzerland . .	29
2.10	Common factors CC: Europe core, Austria and Switzerland . . . .	29
2.11	Common factors CC: Europe core . . . . .	30
3.1	Size of the asymptotic SMS( $\rho_0$ ) test . . . . .	51
3.2	Size of the bootstrap SMS( $\rho_0$ ) test . . . . .	52
3.3	Size distortion and average cycle length . . . . .	55
3.4	Power of the bootstrap SMS( $\rho_0$ ) test against $\rho = 0.10$ . . . . .	56
3.5	Power of the bootstrap SMS( $\rho_0$ ) test against $\rho = 0.25$ . . . . .	57
3.6	Monthly Asian stock index prices: Bulls and bears classification . .	64
4.1	Monthly Asian stock index prices: Bulls and bears classification . .	85
5.1	Returns of stocks, bonds, T-bills, and gold . . . . .	109
5.2	Recursive cross-section test: Univariate . . . . .	113
5.3	Stocks and bonds: Rolling and recursive correlation . . . . .	117

---

## LIST OF FIGURES

5.4	Stocks and T-bills: Rolling and recursive correlation . . . . .	117
5.5	Stocks and gold: Rolling and recursive correlation . . . . .	118
5.6	Bonds and T-bills: Rolling and recursive correlation . . . . .	118
5.7	Bonds and gold: Rolling and recursive correlation . . . . .	119
5.8	T-bills and gold: Rolling and recursive correlation . . . . .	119
5.9	Recursive cross-section test: Co-crashes I . . . . .	127
5.10	Recursive cross-section test: Co-crashes II . . . . .	128
5.11	Recursive cross-section test: Co-crashes III . . . . .	129
5.12	Recursive cross-section test: Flight to quality and liquidity I . . . .	130
5.13	Recursive cross-section test: Flight to quality and liquidity II . . .	131
5.14	Recursive cross-section test: Flight to quality and liquidity III . . .	132
5.15	Recursive cross-section test: Co-crashes and flight to quality/liquidity	133
6.1	S&P500 (partial) autocorrelation . . . . .	157
6.2	S&P500 range and log-range . . . . .	158
6.3	S&P500 range and log-range unconditional distributions . . . . .	158
6.4	Regime probabilities (only the constant changes) . . . . .	164
6.5	Regime probabilities (all mean equation parameters change) . . . .	167
6.6	S&P500 range and smoothed probabilities . . . . .	171
6.7	T-bills range and smoothed probabilities . . . . .	171
6.8	T-notes range and smoothed probabilities . . . . .	172
6.9	FTSE range and smoothed probabilities . . . . .	172
6.10	NIKKEI range and smoothed probabilities . . . . .	173
6.11	Concordances . . . . .	174





# INTRODUCTION

Einstein already realized that global interdependencies are important and too complex to be fully understood and predictable by human beings. More recently also politicians like a former US President became aware that this is true in a socio-economic context, too:

You live in the age of interdependence. Borders don't count for much or stop much, good or bad, anymore. (William Jefferson Clinton<sup>1</sup>.)

Such interdependence is especially notable for financial markets. A former chairman of the New York stock exchange states:

The globalization of the capital market is actually part of economic globalization. This will create a change in the entire world economy, not just restricted to some fields in some countries. (Richard Grasso<sup>2</sup>.)

So, interdependencies among financial markets across the world are strong and possibly increasing in importance. This thesis will therefore be concerned with interdependencies of asset markets and possible changes of these through time.

## 1.1 The general idea of this work

National and international asset/capital market interdependencies or linkages (we will use both terms interchangeably) are strong. Moreover, these linkages have changed in the past. As such, one might first wonder how to measure the strength of such market interdependencies and second how to identify if such linkage-measures have changed in the past. This work highlights some important aspects of these questions. It is an empirical study describing the variation in international market linkages. It does not develop theories explaining such variation, nor is it

---

<sup>1</sup>Harvard University Albert H. Gordon Track And Tennis Center, November 19, 2001

<sup>2</sup>Location and date is unknown

normative. Theoretical contributions focus on econometric testing and modeling procedures in order to better grasp and explain asset market linkages as well as possible changes of those. Further, it proposes different means for measuring such linkages and focuses on national and international asset market prices and their volatilities.

## 1.2 Measures and sources of asset market linkages

Early research on stock market linkages mainly documented cross border return correlations (see e.g. Roll (1988)). Correlation is a static measure of comovement between two or more<sup>3</sup> series like stock market returns or volatilities. It measures the degree of common movement (up or down) over the full sample. A limitation of this basic concept is that it does not take dynamics into account, i.e. current observations do not depend on past information.

Dynamics can be decomposed into short and long run movements. Typically, long run comovement are described by the concept of cointegration developed by Engle and Granger (1987) and is applied to financial markets, for example, in Kasa (1992) and Click and Plummer (2005). If cointegrated, two (or more) time series possess a common (stochastic) trend. In other words, in the long run the levels of these series (asset prices for example) have a common equilibrium, also called common attractor. So, the series will converge to a mutual steady state.

Concerning the short-run, common dynamics would indicate a common behavior of the seasonal and/or business cycle component or the volatility. Short-run common dynamics do not refer to the asset price levels but rather to the comovement of returns (or the first difference of the price levels) and volatilities of the different time series. Typical examples of short-run dynamic comovement measures are common autoregressive conditional heteroscedasticity (ARCH) models (Engle and Susmel (1993)), common features (Engle and Kozicki (1993)), or common factors models (Engle et al. (1990) and Engle et al. (1994)). At this point it is sufficient to think of the degree of market comovements as a measure of similar patterns among series. More exact definitions will be provided in the following chapters.

Many economic phenomena have an important impact on the interplay of financial markets. International economic/financial integration on the one hand and financial contagion on the other hand are probably the two most often suggested reasons for changing comovements of asset markets in both the academic and the popular literature.

---

<sup>3</sup>In such a case we can use partial correlation as in Anderson (1984).

The theory of international economic integration (Balassa (1961)) states that a continuous decline in trade barriers implies higher integration among countries' economies. Such an integration usually follows a slow and gradual process like in the European Union (EU) and German unification. The economic (and often also political) integration process will be accompanied by a financial integration process as well. Financial integration is characterized by lifting controls of capital/financial accounts and the abolishment of other barriers to international investments. This leads to a gradual increase in asset market interdependence among the involved countries. Contagion, on the contrary, is related to the transmission of economic crises from one country to another and is thereby of an abrupt and more short-term nature than economic integration. So, different speeds of changing comovement behavior across asset markets are probably related to different economic phenomena.

Both phenomena can also be connected. For example, countries which are in an economic integration process can be struck by a contagious crisis actually accelerating the integration process, by promoting more political coordination. Examples of such a policy coordination are common money supply changes or the introduction of capital controls. In such a case contagious crises might increase financial comovements in the short-run but also shift the structurally (caused by economic integration) increasing comovement path up. So, a contagious crisis might move up the degree of financial market comovement and economic integration. Another possibility is that after a crisis the degree of comovement falls back to the before-crisis level and continues its gradual increase as determined by economic integration. Below we present both concepts, economic integration and financial contagion, in more detail.

International economic and financial integration tends to strengthen international asset market linkages. For example, Goetzmann et al. (2005) find that international equity correlations change a lot through time showing peaks in the late nineteenth century, the Great Depression, and the late twentieth century. So, mostly periods with free capital flow are associated with high correlations of asset markets. Obstfeld and Taylor (2002) document the relation between integration of global capital flows and relate this directly to the temporal variation in the average correlation of world equity markets.

Chen and Knez (1995) and Korajczyk (1996) test for international market integration using asset-pricing models that consider the international variation in the price of risk. The contributions by Bekaert and Harvey (1995; 2000) provide evidence that market integration and financial liberalization change the comovements of emerging stock market returns with a global market factor. In a newer study Bekaert et al. (2005) find that the US influence has increased on most countries,

suggesting an increased linkage among them. They interpret this as stronger economic integration. So, empirical evidence favors the conclusion that international economic integration has a significant impact on financial market comovements.

But what are the reasons for the increase in market comovements triggered by economic integration? A possible explanation lies in the fundamentals like money supply, foreign reserves, gross domestic product and its growth rate, the current account or the exchange rate system (Calvo and Reinhard (1996), Claessens et al. (2001), and Forbes and Rigobon (2002)). An example is international trade in goods and services (Gerlach and Smets (1995), Chen and Zhang (1997) and Gelos and Sahay (2001)). If two countries are direct trading partners, changes in the economic conditions in one country are directly felt by the other one through changes in demand and supply of traded goods (Gelos and Sahay (2001)). The trade channel can also indirectly affect comovements. When two countries' exports compete in a third country, shocks to that third country would be felt in the two exporting countries at the same time. Furthermore, the financial system itself or multinational organizations can serve as transmission channels as well. For example, internationally operating banks can transfer shocks from one country to another through their balance sheet. Alternatively, foreign direct investment and macroeconomic policy coordination among countries might foster transmission channels. Moreover, common regional or global factors like oil price shocks or military conflicts can also lead to asset market comovement.<sup>4</sup>

How can contagion play a role in international asset market comovements and how can we think of it in the first place? Contagion has been defined in various ways in the literature (Eichengreen et al. (1996), Kaminski and Reinhart (2000), and Edwards and Rigobon (2002)). A very good survey on contagion is the work of Pericoli and Sbracia (2003) who summarize the most representative definitions of contagion. Definitions include: "when volatility of asset prices spills over from the crisis country to other countries", a "sharp fall in the stock market index or with an upsurge in the volatility of asset prices". Other definitions they give are: "a significant increase in comovements of prices and quantities across markets, conditional on a crisis occurring in one market or group of markets", or alternatively a situation "when the transmission channel intensifies or changes after a shock in one market". The last definition was also called "shift-contagion" in the very influential paper by Forbes and Rigobon (2002). Forbes and Rigobon (2002) also distinguished different transmission channels either as being "crisis-contingent" or "non crisis-contingent".

---

<sup>4</sup>Here one might disagree to speak of asset market interdependence because the cause of the comovements are solely factors affecting all markets at the same time but are not necessarily symptoms of linkages among them. We include common factors here for the sake of consistency with the relevant literature.

According to crisis-contingent theories international transmission channels of shocks differ during crisis periods from those during non-crisis times. They include: multiple equilibria based on investors psychology (see, for example, Obstfeld (1996) and Masson (1999)), endogenous-liquidity shocks leading to portfolio recomposition, and political economy which might affect exchange rate regimes. These theories can explain shift-contagion defined above.

Multiple equilibria occur when a crisis in one country serves as a sunspot for other countries which can shift from a good to a bad equilibrium. Such a shock transmission can be caused by investors expectations and not by real linkages. Therefore, the transmission of shocks occurs through channels that do not exist during stable periods. Drying-up of liquidity after the initial shock also leads to increased asset price correlation. The political transmission channel explains why exchange rate crises might occur together through a mechanism triggered by the initial crisis. All these theories share one important feature: transmission mechanisms during crisis periods differ from those before the shock. Theories not having this feature cannot explain shift-contagion and are called non-crisis-contingent theories. Those theories assume that cross-market linkages do not change after a shock. Observed cross-market correlations after shocks are just a continuation of (real) linkages already existing before a crisis hits. Examples of such theories include, among others, international trade, country re-evaluation, policy co-ordination, country similarities and random aggregate shocks.

### 1.3 The importance of asset markets

Global asset markets are central to the economy and serve important purposes. National and international trade in assets provides essential ingredients for efficient economic systems, such as (inter)national risk sharing and intertemporal trade. So, not only investors but also politicians and whole economies have an interest in a proper functioning of asset markets.

Ongoing globalization has contributed to strengthening social, economic and financial ties among countries by increasing international trade and foreign direct investment, the formation of trading blocks and integrated economic systems such as the EU, the emergence of multinational corporations and financial intermediaries, technological and financial innovations, deregulation and market liberalization processes.<sup>5</sup> Such stronger real and financial links between economies usually also lead to more pronounced linkages of their asset markets. Nowadays basically everybody can, within seconds, invest funds in stocks, bonds and all sorts of assets on any asset market in the world. Hereby, news in one part of the world can at the

---

<sup>5</sup>See, for example, Jeon and Chiang (1991), and Chen et al. (2002).

same time influence asset price movements in many other regions of the globe. So in general, asset markets are of central importance for economies. These markets become, at the same time, increasingly interconnected around the world, leading to more and more external factors having an influence on international economies.<sup>6</sup>

Modern portfolio theory introduced by Markowitz (1952) proposes how rational investors should use diversification with imperfectly correlated assets in order to optimize their portfolio with respect to its expected return and its risk or volatility.<sup>7</sup> Investors can diversify their portfolios internationally and/or across assets if this appears to be optimal in an expected risk-return sense. Studying international asset market comovements like correlations of returns and dynamics of volatilities is thereby essential in making such an optimal asset allocation decision.

Apart from its meaning for an optimal portfolio decision, a systematic analysis of asset market linkages is furthermore important because of likely effects of (extreme) asset price movements on real economic activity. Assets are one of the most important factors for people's wealth and thereby for the wellbeing of whole economies. According to basic economic theory, wealth is one of the most important determinants for consumption and investment decisions.

Also changes in the volatility of asset prices can have strong impacts on real economic activity. Increased volatility makes it more difficult for agents to forecast their future wealth levels and might easily distort consumption, savings, and investment decisions. Hence, real economic activity can be greatly affected by swings in assets' prices and a better understanding of such price movements and volatilities might help to cushion such effects.

Furthermore, (negative) asset price changes can have important effects on the stability of financial systems of especially developing but also developed countries. A strong negative shock to an asset market might lead to a deterioration of the asset position of, for example, hedge funds with potential domino effects for leveraged banks and financial systems of entire economies. Examples are the European Monetary System crisis in 1992-93 and the breakdown of the Long-Term Capital Management (LTCM) hedge fund in 1998, which ultimately folded in 2000. Both events triggered uncertainty and strongly volatile asset markets worldwide. In the LTCM crisis, triggered by high leverage and unfavorable asset market movements, only massive national and international bail-out programs by national governmental organizations like the US Federal Reserve Bank (Fed) and private banks could prevent in 1998 the collapse of the fund and possible domino-effects within the fi-

---

<sup>6</sup>As references about the impact of globalization and economic integration on asset markets see, among others, Calvo and Reinhart (1996), Claessens, Dornbusch, and Park (2001), Forbes and Rigobon (2002), Obstfeld and Taylor (2002), and Goetzmann et al. (2005).

<sup>7</sup>For his contributions Markowitz was awarded the Nobel Price of Economics in 1990.

nancial sector with far-reaching economic consequences.<sup>8</sup> A more recent example of events having a global economic impact is the subprime mortgage crises in 2007 and 2008.

When asset markets deliver suboptimal results for national economies, politicians should adapt their policies for these markets. One prominent example of rules by which many national governments have tried to isolate their national asset markets from international influences is capital controls. Academics have not yet reached a consensus on the question if capital controls are welfare enhancing or not.<sup>9</sup> Nevertheless, in general any conclusions on the usefulness of certain policies hinge on a detailed knowledge of the functioning, dynamics, and linkages of international financial markets. For example, negative news in one country or region can cause investors to withdraw funds from initially unaffected markets leading to negative repercussions for these countries as well. Politicians should then strive to insulate their national asset markets from international ones.

## 1.4 Aim of the study

Research on asset markets and their dynamics and interdependencies has made considerable progress in the last decennia. Nevertheless, there is still scope for advancement in this field because important issues like cyclicity of asset markets have not sufficiently been dealt with yet. Furthermore, financial and governmental institutions, but also investors, ask for more and more sophisticated tools to model financial markets and their comovements as a basis for policy, investment, or pricing decisions.

Comovements of asset markets can have many forms. This thesis provides a systematic approach towards a range of these possible dimensions of asset market comovements. We not only have a look at the comovements of asset price returns, but propose a number of different approaches. Such approaches include co-jumps of asset prices, the cyclical co-occurrence of bull and bear phases, co-extreme returns or tail-dependency opposed to return correlations calculated over the full sample, and the comovement of the volatility of asset returns as well as the co-occurrence of volatility regimes across markets. We also consider the regional dimension (especially East-Asia) versus global comovements. Overall, there is clear evidence for an increase in comovements of international asset markets' price processes and volatilities especially in the last ten to fifteen years.

The aim of this thesis is two-fold. First, we aspire to contribute to the aca-

---

<sup>8</sup>See Lowenstein (2000), Dunbar (2001), MacKenzie (2003), and Fenton-O'Creevy et al. (2004).

<sup>9</sup>See, for example, De Gregorio et al. (2000).



demic discussion on how certain characteristics of international asset markets are linked and how these links and comovements have developed over time. Among others, we propose an innovative approach to the problem of changing correlations across asset markets. Up to now, most academic work has focused on raw returns with either daily, weekly or monthly data frequencies, which usually biases pure correlation measurement in the presence of heteroscedasticity. We rather build on the idea of extracting and analyzing the cyclical component of returns instead of the raw returns themselves. We show that investigating any changes in comovements of these components is not biased by heteroscedasticity in the raw data. This innovative approach of looking at the cyclical component of asset markets adds a new dimension to the existing knowledge about asset market comovements and can be important for investors with a cyclical strategy or a medium-run investment horizon.

The second main aim of this thesis is to help enlarging our understanding of the functioning of asset markets and also improve our forecasting capabilities for asset markets' volatility by introducing new econometric tools. We propose a new time series model for the estimation and forecasting of volatility, which then will be applied to asset market data in order to check for comovements in markets' volatilities. We show that this model performs better than mainstream GARCH-models. Furthermore, we analyze the comovements of the realized volatility and co-jumps of some industrialized countries' stock markets.

This work falls short of definitely answering the question of what exactly has caused the increase in asset market comovements that we find. If these increased comovements are driven by economic integration or only by contagious crises remains an open question. We actually find evidence for both explanations depending on the markets and samples we look at. Nevertheless, time series might still be too short today in order to clearly judge in favor of either hypothesis. We therefore think that future research should focus on methods that are able to distinguish between financial integration and contagion as sources for changing international asset market comovements. A good starting point is the work of Candelon and Bodart (2004) who propose a frequency domain approach in order to disentangle the more contagious short-run and the long run effects coming from financial integration. Nevertheless, they only consider the possibility of causal effects between integration and contagion. One causal relation from interdependence to contagion is obvious as stronger financial and economic linkages lead to higher vulnerability towards asymmetric shocks. But it may also be the case that contagion induces higher integration because of common political reactions of countries towards solving a crisis.

## 1.5 Structure of the thesis

This section concludes the first chapter and introduction of this thesis by outlining its structure and shortly summarizing every individual chapter. Basically all chapters, excluding Chapter 1, correspond to one essay/paper each. The original papers are mainly identical to the chapters. As the chapters/essays differ somewhat with respect to applications, econometric methodologies and focus, we refrain from placing them into the literature at this point. We rather get back to this in each chapter separately. In total, the thesis consists of six chapters. We refrain from writing an overall conclusion in a separate chapter. In our opinion it would not add much to the conclusions in the chapters themselves as the chapters are identical to the underlying papers.

Chapter 2 analyzes common factors in the continuous volatility component, co-extreme and co-jump behavior of a sample of stock market indices. In order to identify those components in stock price processes during a trading day we use high-frequency data and techniques. We show that in most of the cases one common factor is enough to describe the largest part of the international variation in the continuous part of volatility and that this factor's importance has increased over time. Furthermore, we find strong evidence for asymmetries between extremely negative and positive co-extreme close-open returns and of negative and positive co-jumps across countries. Thereby, Chapter 2 also highlights routes followed in the remaining chapters of the thesis.

Chapter 3 implements estimation and testing procedures for comovements of stock market “cycles” or “phases” in Asia. We extend the Generalized Method of Moments (GMM) approach to measuring business cycle synchronization due to Harding and Pagan (2006b) toward estimating and testing for (bivariate, multivariate) cyclical stock market synchronization in a small sample framework. We show that the asymptotic version of the test gets increasingly distorted in small samples when the number of countries grows large. However, a block bootstrap version of the test can remedy the size distortion when the time series length divided by the number of countries  $T/n$  is sufficiently large. Applying the financial cycle synchronization framework to five East-Asian economies, we are unable to reject the null hypothesis of a nonzero common multivariate synchronization index for certain economically meaningful subsets of these countries.

Chapter 4 extends the ideas of Chapter 3 and proposes a test for endogenously determining structural change in the bivariate and multivariate synchronization indices. Upon applying the technique to five Asian stock markets we find a significant increase in the cross country comovements of Asian bullish and bearish periods in 1997. A power study of the stability test suggests that the detected increase in comovement is more of a sudden nature (i.e. contagion or “Asian Flu”)

instead of gradual (i.e. financial integration). It is furthermore argued that stock market cycles and their propensity toward (increased) synchronization contain useful information for investors, policy makers as well as financial regulators.

Chapter 5 assesses the linkages between the most important U.S. financial asset classes (stocks, bonds, T-bills and gold) during periods of financial turmoil. Our results have potentially important implications for strategic asset allocation and pension fund management. We use multivariate extreme value theory to estimate the exposure of one asset class to extreme movements in the other asset classes. By applying structural break tests to those measures we study to what extent linkages in extreme asset returns and volatilities are changing over time. Univariate results and bivariate comovement results exhibit significant breaks in the 1970s and 1980s corresponding to the turbulent times of e.g. the oil shocks, Volcker's presidency of the Fed or the stock market crash of 1987.

Chapter 6 proposes a new model to describe asset market volatilities. It suggests to use the range, defined as the percentage difference between the maximum and the minimum price of an asset, as a nonparametric proxy for an asset's volatility. We find that range time series can very well be modeled in a Markov-switching (MS) autoregressive moving-average (ARMA) with generalized autoregressive conditional heteroscedasticity (GARCH) framework. Such a model allows to identify different volatility regimes, i.e. a high and a low volatility regime, and to forecast volatility  $h$  periods into the future. In an international comparison we find evidence for volatility comovements across different countries.

Here we would like to point out that the notation with respect to, for example, variables and time series is not always consistent and may change slightly from chapter to chapter. Nevertheless do we stick to the same notation within every chapter itself.

# COMOVEMENTS OF RETURNS AND VOLATILITY IN INTERNATIONAL STOCK MARKETS: A HIGH-FREQUENCY APPROACH

This chapter analyzes common factors in the continuous volatility component, co-extreme and co-jump behavior of a sample of stock market indices. In order to identify those components in stock price processes during a trading day we use high-frequency data and techniques. We show that in most of the cases one common factor is enough to describe the largest part of the international variation in the continuous part of volatility and that this factor's importance has increased over time. Furthermore, we find strong evidence for asymmetries between extremely negative and positive co-extreme close-open returns and of negative and positive co-jumps across countries. Thereby, this chapter also highlights routes followed in the remaining chapters of the thesis.

## 2.1 Introduction

Financial economists have extensively documented so-called “empirical stylized facts” of stock market returns such as clusters of volatility and heavy tails, see e.g. Embrechts et al. (1997). Also comovements or synchronization of asset markets have been analyzed. Early research on stock market linkages mainly documented cross border return correlations (see e.g. Jaffe and Westerfield (1985) and Roll (1988)). Contributions like Lin et al. (1994) and Susmel and Engle (1994) used ARCH-type models to investigate the direction of international spillovers as well as to identify differences in market comovements during periods of market tur-

bulence and market quiescence.<sup>1</sup> A relatively new strand in the literature uses high-frequency returns in order to extend the information set to analyze financial markets and their comovements. Examples include Andersen et al. (2007), Barndorff-Nielsen and Shepard (2001; 2003; 2004).

Using high frequency data has a number of advantages. First of all, having more rather than less data can never be a disadvantage. It allows a more accurate analysis of, for example, the volatility of the underlying data generating processes. Daily, weekly or monthly data may be too aggregated in order to find important patterns in the data. Especially financial markets are known to react immediately to news rather than after a day or a week.<sup>2</sup> Without high-frequency data, intradaily patterns of stock returns could not be examined. Second, as in Andersen et al. (2007), Barndorff-Nielsen and Shepard (2001; 2003; 2004), a common assumption is that the logarithmic stock price follows a continuous-time jump diffusion process with a continuous and a jump component to it. Using high-frequency data allows us to consistently estimate the integrated volatility in a continuous time diffusion model and gives asymptotically much more accurate volatility estimates than models based on low-frequency data.

In this chapter we would like to analyze, among others, the continuous component and the jump behavior of some industrialized countries' stock markets. In order to identify those components in stock price processes during a trading day we have to use high-frequency data. Thereby, we extend the literature about common factors in volatility, like in Engle and Susmel (1993), and also look at co-jumps across markets. Co-jumps have not been treated much in the literature with notable exceptions like Gobbi and Mancini (2006) and Lahaye et al. (2007).

The main objective of this chapter will be to identify common factors in the continuous component of volatility and co-jumps across eight industrialized countries. By using the continuous component we would like to determine how much variation in the data is common and how this has changed over the sampling period. Common factor models are widely used in financial applications. Theoretical models like the Arbitrage Pricing Theory, which was introduced by Ross (1976), allow multiple risk factors to determine assets' return dynamics. Factors can include interest rates, GDP growth rates, investment, and other macroeconomic variables. As macroeconomic conditions across countries develop similarly, common factors could be observed. Such theory has also led to volatility factor models as for example in Diebold and Nerlove (1989), Engle et al. (1990), and Engle et al.

---

<sup>1</sup>Baillie and Bollerslev (1989), however, argued that the modelling of returns can result in a loss of information on possible common trends when prices are cointegrated. Researchers have therefore also turned attention to VAR frameworks that take cointegration and error correction into account. Representative articles of the cointegrated VAR literature are Ghosh et al. (1999), Chen et al. (2002), and Click and Plummer (2005).

<sup>2</sup>See, for example, Beine et al. (2007) and Beine et al. (2007).

(1994). Such models specify the conditional variance parametrically, which can make estimation difficult especially when the amount of time series increases.

We also test if there are asymmetries in the co-jump behavior. In other words, we would like to test if positive and negative co-jumps behave differently across countries. Further, we will analyze possible dependencies between extreme close to open returns and the jump behavior on the following trading day. As Andersen et al. (2007) already pointed out, the close to open or overnight returns account for a non-negligible part of the total return volatility. They find that, for example, for the S&P500 and the T-Bond markets, the overnight returns account for roughly sixteen percent of the total volatility. As such we think that we would miss potentially important information by not taking those returns and possible synchronization across markets into account.

Investors and speculators who follow real time trading strategies are interested in high-frequency interrelations of asset markets to optimally time their portfolio rebalancing. For some early contributions to the theory about optimal portfolio choice models like the capital asset pricing models see, among others, Markowitz (1952) and Sharpe (1964). Also, extreme movements in stock market prices and the potentially destabilizing effects on the real economy raise the issue of how monetary authorities should respond. Indeed, bullish stock markets can induce large amounts of loan collateral which then increase demand and goods price inflation. Moreover, when the stock market decreases rapidly, this can result in widespread liquidity problems and a “credit crunch” in the financial system. Thus, monitoring the impact of stock market swings and volatility is also of potential interest to regulatory bodies caring about systemic risk and overall financial stability. Finally, if stock markets have become more synchronized over time, the potential for financial system instability to spill over to other countries might increase, which would suggest a coordinated effort of policymakers and regulatory bodies.

This chapter therefore highlights important aspects of stock markets’ returns and volatility based upon high-frequency techniques. Furthermore, we combine those techniques with a rich international high-frequency stock index data-set in order to stress characteristics of international market linkages. We would like to use this as an introduction to topics dealt with in later chapters of this work, where we also analyze asset market comovements.

The remaining sections are organized as follows. Section 2.2 discusses the statistical methodology that will be implemented. In Section 2.3 we show the empirical estimation and testing results. Concluding remarks are contained in Section 2.4.

## 2.2 Methodology

### 2.2.1 The basic setup

In this chapter, as is common in the literature, we assume that logarithmic stock prices  $p(t)$  follow a continuous-time jump-diffusion process defined by the stochastic differential equation given by:

$$dp_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t, \quad (2.1)$$

where  $\mu_t$  corresponds to a continuous and locally bounded variation drift process,  $\sigma_t > 0$  denotes the spot volatility process with a right-continuous sample-path with well defined limits and  $W(t)$  is a standard Brownian motion.<sup>3</sup> These first two terms correspond to the continuous part of the total variation process.  $\kappa_t dq_t$  refers to the jump component of the total process, where  $q_t$  is a counting process with possibly time-varying intensity  $\lambda_t$  meaning that  $dq_t = 1$  when a jump occurred at  $t = s$ .  $\kappa_t$  stands for a possibly time-varying size of the jumps. So, the quadratic variation (QV) for the cumulative return process is the integrated volatility of the continuous path component plus the sum of the  $q_t$  jumps that have occurred between  $t = 0$  and  $t = t$ . This can be written as:

$$QV_t = \int_0^t \sigma_s^2 ds + \sum_0^t \kappa_s^2. \quad (2.2)$$

Here again the first integrated variance term measures the contribution from the continuous sample-path variation, while the latter part sums over all squared discontinuities or jumps that occurred until time  $t$ .

Quadratic variation or volatilities in general can be evaluated at any frequency the researcher likes. Mostly, daily, weekly, or monthly frequencies are used. In this paper we focus on the daily volatility. In order to obtain those we need a measure for intradaily returns, which will be given by:

$$r_{t,j} = p\left(t - 1 + \frac{j}{M}\right) - p\left(t - 1 + \frac{j-1}{M}\right) \quad j = 1, 2, \dots, M.$$

Here,  $M$  refers to the amount of equally spaced return observations over one trading day. So, from now on the first part on the right hand side of (2.2) is referred to the continuous daily volatility component and the second part accounts for the contribution from within-day jumps or the discontinuous part of the the

---

<sup>3</sup>See, for example, Barndorff-Nielsen and Shepard (2001; 2003; 2004) and Andersen et al. (2007) among others.

daily variation.

### 2.2.2 Realized measures and jump test statistic

In practice the continuous and the discontinuous part are not observable and therefore need to be estimated by approximation. Having defined the intraday returns we can proceed to the definitions of the realized measures in order to obtain estimates for the continuous and the discontinuous (jump) parts of the processes. We will do this by means of the non-parametric measures of realized volatility (RV) and realized bi-power variation (BPV) advocated in, for example, Andersen et al. (2001; 2003) and Barndorff-Nielsen and Shepard (2004). According to those authors realized volatility can be defined by:

$$RV_t(M) \equiv \sum_{j=1}^M r_{t,j}^2. \quad (2.3)$$

In other words, the realized volatility for day  $t$  is nothing else then the sum of  $M$  available squared intraday returns. As noted in Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shepard (2004)  $RV_t(M)$ , converges in probability to the increment of the quadratic variation process as  $M \rightarrow \infty$ . We can re-write this as:

$$RV_t(M) \xrightarrow{p} \int_{t-1}^t \sigma_s^2 ds + \sum_{s=1}^t \kappa_s^2, \quad \text{for } M \rightarrow \infty. \quad (2.4)$$

This means that in the absence of jumps, RV was a consistent estimator for the integrated volatility.

It is by now agreed, though, that jumps in asset prices are a quite common and frequent phenomenon.<sup>4</sup> Taking this as a starting point Barndorff-Nielsen and Shepard (2004; 2006) introduced another volatility measure, which they called bi-power-variation or just BPV, which is very closely related to RV. BPV now can be used in order to disentangle both the continuous and the jump component from the realized volatility, because it estimates the integrated volatility consistently even in the presence of jumps. Barndorff-Nielsen and Shepard propose, instead of squaring intraday returns, to multiply adjacent absolute intradaily returns and standardizing them by a constant. Only in the case of no jumps BPV is asymptotically slightly less efficient than RV in estimating integrated variance. Bi-power

---

<sup>4</sup>Giot et al. (2007) use data of 100 individual stocks and analyze in how far trading activity influences the continuous and the jump component of volatility. They find that trading activity relates positively to the continuous volatility component but negatively to the jump component. One of their interpretations is that poor trading volume leads to more erratic volatility changes. For details see Giot et al. (2007).



variation can therefore be written like:

$$BPV_t \equiv \mu_1^{-2} \sum_2^M |r_{t,j-1}| |r_{t,j}|, \quad (2.5)$$

where,

$$\mu_1 = E(|Z|^1) = \sqrt{2/\pi}, \quad \text{and} \quad Z \sim N(0, 1).$$

Barndorff-Nielsen and Shepard (2004) prove that  $BPV_t$  consistently estimates the integrated variance when the sampling frequency goes to infinity

$$BPV_t(M) \rightarrow \int_{t-1}^t \sigma_s^2 ds, \quad \text{for} \quad M \rightarrow \infty. \quad (2.6)$$

Autocorrelation in the intraday returns might even at a 5-minute frequency still be a problem and led to the proposal to stagger the absolute returns used in (2.5).<sup>5</sup> Such an approach gives:

$$BPV_t(M) \equiv \mu_1^{-2} \left( \frac{M}{M-2} \right) \sum_2^M |r_{t,j-2}| |r_{t,j}|, \quad (2.7)$$

which is the measure for  $BPV$  we actually implement. So, consequently as  $M \rightarrow \infty$  we can use  $RV$  in conjunction with  $BPV$  in order to estimate contribution of the jumps to the quadratic variation process by just subtracting  $BPV$  from  $RV$  as follows:

$$RV_t(M) - BPV_t(M) \rightarrow \sum_{s=1}^t \kappa_s^2, \quad \text{for} \quad M \rightarrow \infty.$$

In finite samples and a sampling frequency less than infinite we might have negative estimates of the jump component, which is not possible from a theory point of view. A common procedure here is to truncate the jump measure at zero giving us the following rule that we apply:

$$J_t(M) \equiv \max[RV_t(M) - BPV_t(M), 0]. \quad (2.8)$$

Such a definition of jumps might yield very small jumps, which is against the intuition that jumps or discontinuities should be quiet noticeable events in the evolution of asset prices. One may therefore only focus upon statistically significant jumps and attribute the non-significant jumps back to the continuous variation part. To do so, we need a test statistic with which we can distinguish between significant and in-significant jumps.

---

<sup>5</sup>See, for example, Andersen, Andersen et al. (2007) or Beine et al. (2007).

Following Huang and Tauchen (2005) and Andersen et al. (2006) we implement the test statistic, where for ease of notation we suppress its dependency on the amount of intradaily returns  $M$ :

$$Z_t \equiv \frac{\frac{RV_t - BPV_t}{RV_t}}{\sqrt{\left(\left(\frac{\pi}{2}\right)^2 + \pi - 5\right) \frac{1}{M} \max\left(1, \frac{RTQ_t}{BPV_t^2}\right)}}, \quad (2.9)$$

where

$$RTQ_t \equiv M \mu_{4/3}^{-3} \sum_3^M |r_{t,j-2}|^{4/3} |r_{t,j-1}|^{4/3} |r_{t,j}|^{4/3}, \quad (2.10)$$

with  $\mu_{4/3} \equiv 2^{2/3} \Gamma(7/6) \Gamma(1/2)^{-1}$ , and where  $\Gamma$  stands for the Gamma-function. Huang and Tauchen (2005) show that the  $Z_t$  statistic defined in equation (2.9) is asymptotically standard normally distributed under the null hypothesis of no within-period (here a day) jumps and has reasonable power against many plausible stochastic volatility jump diffusion models.<sup>6</sup>

In practice one has to chose a significance level  $\alpha$  in order to set a cutoff for jumps being considered either significant or non-significant. So, the actual time series of only significant jumps is defined as:

$$J_{t,\alpha} = I[Z_t > \Phi_\alpha] \cdot [RV_t - BPV_t], \quad (2.11)$$

where  $I$  refers to the indicator function being equal to one when the argument is true and zero otherwise.  $\Phi_\alpha$  stands for the critical value of the standard normal distribution for a chosen significance level  $\alpha$ . In general, a larger  $\alpha$  means that one considers larger jumps and henceforth less numerous discontinuities in the stochastic process.

Having identified the significant jumps the remainder will be considered the continuous portion of the asset price process, which can easily be summarized in the following equation:

$$C_{t,\alpha} = I[Z_t \leq \Phi_\alpha] \cdot RV_t + I[Z_t > \Phi_\alpha] \cdot BPV_t, \quad (2.12)$$

which automatically ensures that the non-parametric measures for the continuous and the jump component add up to the realized volatility  $RV_t$  as claimed in equation (2.4).

---

<sup>6</sup>In the paper by Huang and Tauchen (2005) they report extensive evidence by simulation showing that the  $Z_t$  test statistic has very good size and power properties for a one-factor logarithmic stochastic volatility plus compound poisson jump process.

### 2.2.3 Common factors in the continuous component

Turning attention to more than one country or multiple markets within the same country, intuition predicts that assets markets' movements are driven by some common factors.<sup>7</sup> In support of this Engle and Marcucci (2006, p.8) argue that:

In finance, there is a strong belief that movements in the price of one particular asset are quite likely to coincide with movements in the prices of other assets, possibly quoted in different markets. These comovements might be caused by the reaction of economic agents to particular changes in some macroeconomic and financial variables or, maybe, to specific news about the company or about the economic sectors involved. In addition, the movements in one asset price may have implications that are likely to affect the value of other assets...

Once we concentrate on countries located in the same geographical region and with similar degrees of economic developments and structures, the common factor hypothesis becomes even more important. Basic economic reasoning might predict that asset markets in a specific region are driven by at least one global economic factor and/or by a regional factor, specific to the region under consideration. This does not preclude any additional common factors.<sup>8</sup>

In a recent paper Anderson and Vahid (2007) argue for principal component analysis as a technique to identify and isolate common market factors in a realized volatility setting. Such motivation in the approximate factor literature in finance goes back to work of Chamberlain and Rothschild (1983). A factor model can be represented as:

$$C_t = \Lambda F_t + u_t, \quad (2.13)$$

where  $C_t$  is a  $(N \times 1)$  vector of, for simplicity, demeaned square roots of the continuous parts of the  $N$  considered price processes,  $F_t$  being a  $(r \times 1)$  vector of  $r$  common factors with factor loadings summarized in the  $(N \times r)$  matrix  $\Lambda$ , and  $u_t$  being a  $(N \times 1)$  vector of  $N$  idiosyncratic disturbances orthogonal to the common factors in  $F_t$ . So, here  $F_t$  contains common second moments or volatility factors whereas Chamberlain and Rothschild (1983) considered first moments or returns. Such multivariate realized volatility factor models are discussed in Lo and Wang (2000), Andersen et al. (2001), and Anderson and Vahid (2007). As noted in Anderson and Vahid (2007), Chamberlain and Rothschild (1983) show that the

---

<sup>7</sup>See for example Chapter 6 in Campbell et al. (1997), and Chapter 9 in Cochrane (2001) for theoretical and empirical considerations about this point.

<sup>8</sup>Another interesting way to study volatility spillovers across markets is the methodology proposed by Cheung and Ng (1996). They develop a test for causality in variance based on the residual cross-correlation function. Such a method is limited to the bivariate case though and we also want to study the common component of more than two time series.

$r$  largest eigenvalues of the matrix  $\frac{1}{T} \sum_{t=1}^T C_t C_t'$  will go to infinity as  $N$  and  $T$  go to infinity, whereas the  $(r+1)^{th}, (r+2)^{th}, \dots, N^{th}$  eigenvalues remain bounded.

In practice with finite samples the researcher has to apply a certain rule for significance of common factors as opposed to idiosyncratic factors. Bai and Ng (2002) develop different model selection criteria for choosing the number of common factors, which depend on an arbitrary choice of  $r^{max}$  being the largest possible number of common factors considered by the researcher. A well known problem with these information criteria is that they tend to chose  $r^{max}$  as the number of common factors in finite samples.<sup>9</sup> We experimented with these information criteria and all of them chose  $r^{max}$  as the amount of common factors, which is not reasonable. So, we decided to add a more intuitive decision rule on top of the Bai and Ng (2002) criteria that identifies factors as common if their corresponding normalized eigenvalue is  $\geq 1$ . This means that factors are considered common if they explain at least as much of the variation in the volatility as an average factor would do.

Anderson and Vahid (2007) further show how outliers in the data can severely distort the principal component estimator for common factors especially when  $N$ , the amount of countries in our case, is small. As can be seen in Figures 2.1 to 2.8 the continuous volatility component shows some large fluctuations, which depend upon the chosen  $\alpha$  level for significant jumps. A larger choice for  $\alpha$  means that less jumps are identified to be significant, hence leading to a more erratic continuous volatility part. Anderson and Vahid (2007) propose to account for these extremes in the continuous part by using an instrumental variable approach in order to alleviate their effects on the factor analysis. Applying their method essentially adds up to only considering that part of the continuous component,  $\int_t^{t+1} \sigma^2(s) ds$ , that is predictable by its own past. So, they propose to use  $C_{t-1}$  as an instrument for  $C_t$ . One might thereby write the instrumental variable factor model as:

$$C_t = \Upsilon C F_t + C J_t + J_t + u_t, \quad (2.14)$$

where  $C_t$  and  $u_t$  are defined as before. Now  $C F_t$  represents the “continuous” common factors with  $\Upsilon$  as a factor loading matrix,  $J_t$  are the jumps identified by the  $Z_t$  statistic, which are assumed to be unpredictable from the past, and  $C J_t$  is the “continuous jump or outlier” part which can also be considered as a residual of the instrumental variable regression. We assume that all regularity conditions on the factor loadings  $\Upsilon$ , cross sectional and time series dependence of common factors and idiosyncratic terms stated by Bai and Ng (2002) for consistency of the principal component estimator of common factors as  $\min(N, T) \rightarrow \infty$ , are

<sup>9</sup>See, for example, Anderson and Vahid (2007).

satisfied. The augmentation of the model by the outlier and the jump parts does not change the asymptotic properties of the principal component estimator because  $CJ_J + J_t + u_t$  still satisfies the necessary condition of idiosyncratic components to  $CF_t$ . We, therefore, use proposition 1 from Anderson and Vahid (2007), which reads as follows:<sup>10</sup>

**Proposition 1:** Under the assumption that  $E(CF_t[C'_{t-1}, \dots, C'_{t-p}])$  has rank  $r$ , a consistent estimator of common factors as  $N, T \rightarrow \infty$  with  $N < T$  is  $\hat{\Upsilon}'_{IV}C_t$ , where  $\hat{\Upsilon}_{IV}$  consists of the eigenvectors corresponding to the  $r$  largest eigenvalues of  $\hat{C}\hat{Y}'$  and  $\hat{C}$  is the orthogonal projection of  $C$  on  $C_{-1}$ . Here,  $C = (C_{p+1}, \dots, C_T)$  is  $N \times (T - p)$  and  $C_{-1}$  is the  $Np \times (T - p)$  matrix of lagged values, i.e.

$$C_{-1} = \begin{pmatrix} C_p & & C_{T-1} \\ \vdots & \dots & \vdots \\ C_1 & & C_{T-p} \end{pmatrix}$$

for some  $p > 0$ . Subject to the normalization that  $\Lambda'\Lambda = I_r$ , this estimator is also the ordinary least squares reduced rank regression estimator of  $\Lambda$  in

$$Y = \Lambda BY_{-1} + U$$

that minimizes  $tr(UU')$ .<sup>11</sup> For the proof of this proposition see Lütkepohl (1991) and Anderson and Vahid (2007). For simplicity and as suggested in Anderson and Vahid (2007) we assume  $p = 1$  here.

### 2.2.4 Co-jumps

Obviously, it is not only likely that countries' asset markets are driven by common factors in the continuous component of volatility but that especially discontinuities or jumps in asset prices are transmitted across countries' borders one way or another. Therefore, it is also very interesting to check for co-jumps, which have to account for possible global or regional common jump-factors.

**Definition co-jumps:** Two or more countries' asset markets are said to co-jump if their univariate jump test statistics  $Z_t$  defined in (2.9) are significant on the same day  $t$  and the sums of the largest intra-day returns contributing to the significance of the test statistic  $Z_t$  on that day  $t$  have the same sign.

In order to identify possibly more than one intra-day return that contribute to such a significance we will follow an intuitive procedure already proposed in, for example, Beine et al. (2007). For each day where we find a significant jump test

---

<sup>10</sup>Where Anderson and Vahid (2007) use a slightly different notation as we do.

<sup>11</sup>Here  $I_r$  is the  $r$ -dimensional identity matrix and  $tr$  is the trace operator.

statistic, we first take the largest absolute intra-day return. This return is then set equal to zero and we recalculate the jump test statistic for that day all over again. If it is still significant we also set the second largest absolute intra-day return equal to zero and again redo the whole procedure until the jump test statistic is not statistically significant anymore at the chosen significance level. Like this we obtain all intra-day returns defined as jumps according to the definition. Such returns can either all be positive, negative or both, depending on the underlying causes for the jumps. In order to draw conclusions about the direction the jump component influences the price process, we take the daily sum of all identified intra-day returns, which constitute jumps. Logically, such a sum can either be positive or negative.<sup>12</sup> According to the sign of the sum of intraday jumps we can group days showing jumps into a “negative” and a “positive” jump component. Only days where the sums of the jumps have the same sign across countries qualify to be considered as co-jumps, because these are assumed to be caused by some common regional or global factor which affects all countries’ asset markets in the same way.

## 2.3 Empirical results

In this section we want to establish “stylized facts” and characteristics of stock markets and focus on the main markets in Europe and partly the US. In particular we have five minutes stock market data for Austria, France, Germany, Italy, the Netherlands, UK, US, and Switzerland all starting on January 2, 1997 and ending on June 30, 2006. For all countries we obtained the data for the main stock market index in each country. We will apply the methods introduced in Section 2.2 of this chapter. Our focus is going to be the decomposition of the individual stock market volatilities into a continuous and jump component and close-open returns. Further we will examine comovements of these measures across countries’ markets. Such an analysis will allow us to draw conclusions about possible interconnections of asset markets and will serve as an introduction to the analysis in the following chapters of this thesis.

### 2.3.1 Continuous and jump component

As outlined in Section 2.2 we can decompose the stock price processes in our dataset into a continuous and a jump component. In Figures 2.1 to 2.8 we plot the constructed series for the continuous ( $C_\alpha$ ) and the jump component ( $J_\alpha$ ) of the

---

<sup>12</sup>Obviously, zero would theoretically also be possible but is very unlikely and never occurred in the data set.

stock markets of all eight countries in our sample. The upper panel shows the continuous component whereas the lower one shows the jump component. Both series have been derived by using a 99.9% critical value for the test statistic in Equation (2.9).<sup>13</sup> Many series show, for example, a very large jump on 9/11/2001, the day of the terrorist attacks in New York. Other periods of increased volatility and jump activity are around the Asian crisis (1997), the Russian crisis in conjunction with the LTCM hedge fund collapse (1998), and the Brazilian crisis (1999). Furthermore, the period from mid 2002 until mid 2003 displays strongly increased volatility in many countries in the sample, which might be attributed to the uncertainty in world financial markets before and during the second Gulf War, which started in March 2003. In Chapters 5 and 6 we revisit these periods of increased volatility in more detail.

In Figures 2.1 to 2.8 one can also see that the average sizes and amounts of jumps vary across countries. Therefore, we summarize some descriptive statistics of the continuous and the jump component series in Table 2.1 for all the countries. The means of the jump components differ a lot across countries. For example, the Netherlands show an average of 0.69 whereas Austria only displays an average of 0.27.<sup>14</sup> Another interesting feature of the jump component is that the relative number of trading days during which we observe a jump at the chosen significance level  $\alpha$  differs very much across countries as well. Again Austria stands out with a very high unconditional probability of 17.9% for observing a significant jump on any trading day. Whereas in the US only 4.4% of the trading days show a jump. Assuming binomial distributed jump occurrences, such a difference is very significant. In general, the jump component can be characterized as strongly non-normal, with Jarque-Bera p-values all being significantly below 1%, strongly right-skewed and showing excess Kurtosis.

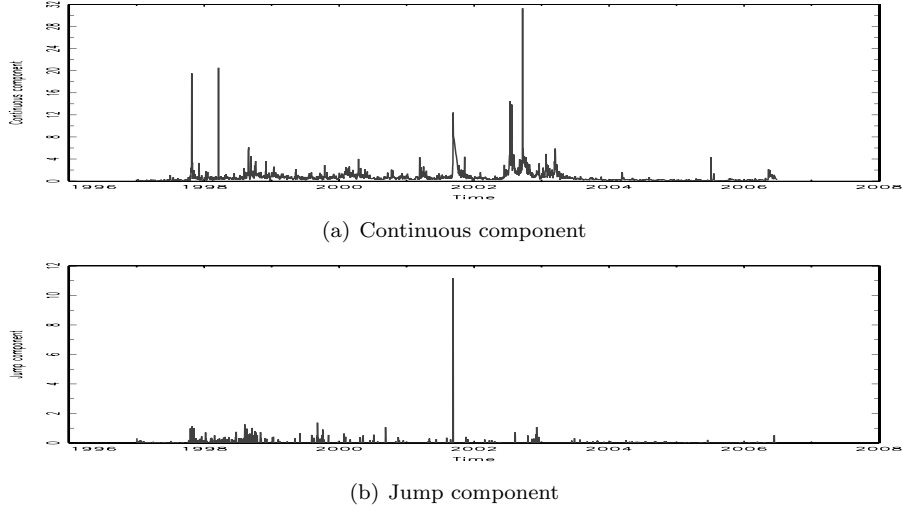
Also the continuous component series are summarized in Table 2.1. Similar to the jump component results, one can see much variance in the measures of centrality (mean and median) of the unconditional distributions. The lowest median has Switzerland with 0.30 which compares to the largest one from Germany having 1.05. Also the variation in the continuous component observations differs very much across countries, which might be deduced from the standard errors. Lastly, by observing the clear rejection of the null hypothesis of the Dickey-Fuller

---

<sup>13</sup>In the literature on realized volatility different critical values have been used. See, for example, Lanne (2007). Obviously increasing  $\alpha$  reduces the amount of significant jumps and increases some of the continuous part observations where the jump would become insignificant. Later in this chapter we also experiment with a significance level of  $\alpha = 99.99\%$ .

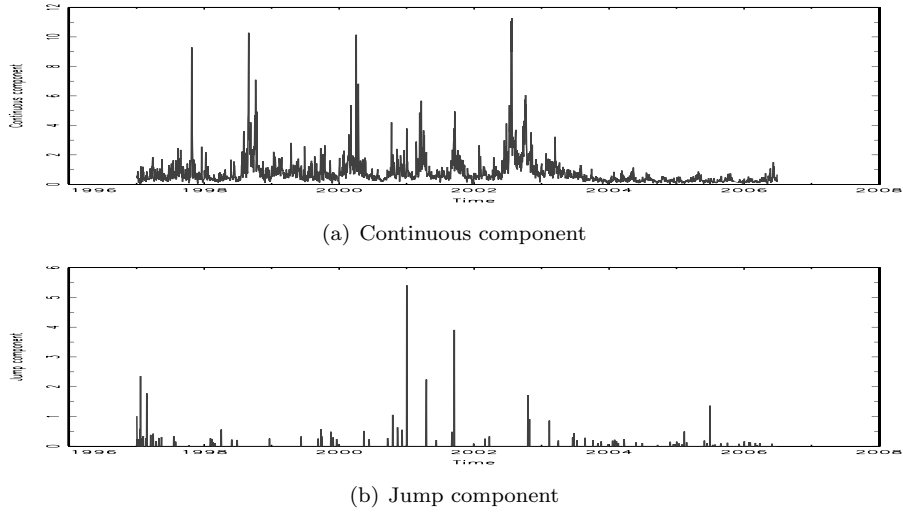
<sup>14</sup>We do not statistically test for the significance of the differences, though, because one would have to make assumptions about the distribution of the size of the jump component which is generally unknown. Here, we only want to highlight that there is quite some variation in the means of the series.

Figure 2.1: UK: Continuous and jump component



**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

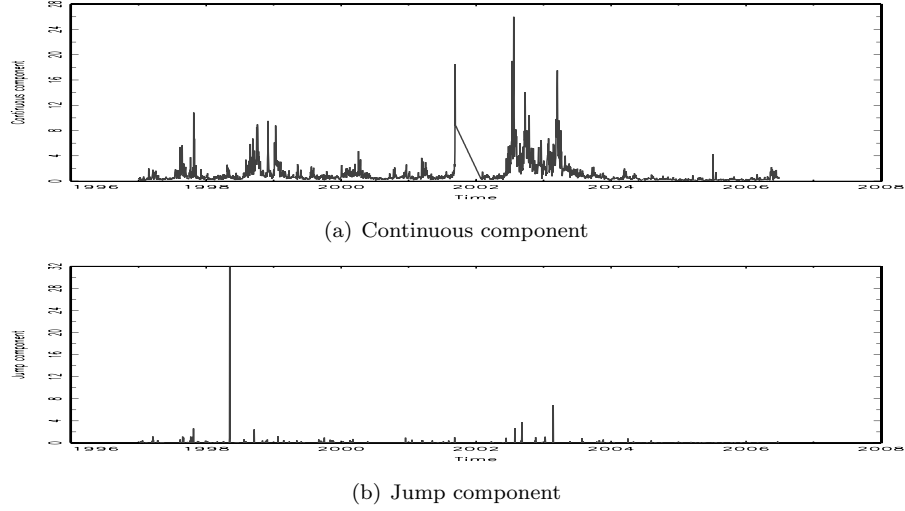
Figure 2.2: US: Continuous and jump component



**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

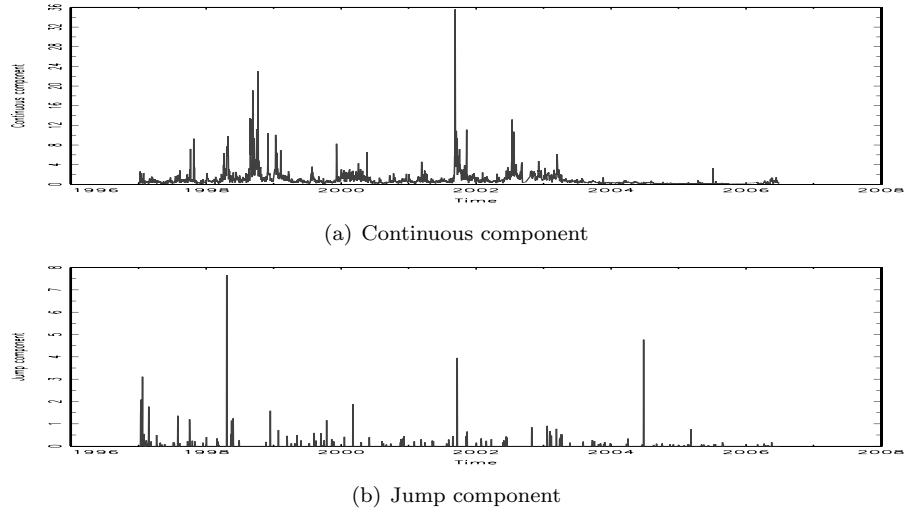


Figure 2.3: The Netherlands: Continuous and jump component



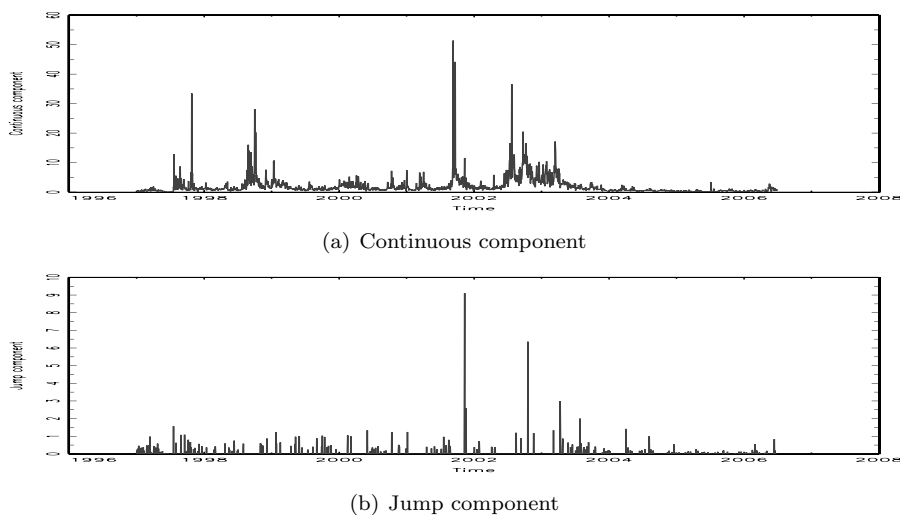
**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

Figure 2.4: Italy: Continuous and jump component



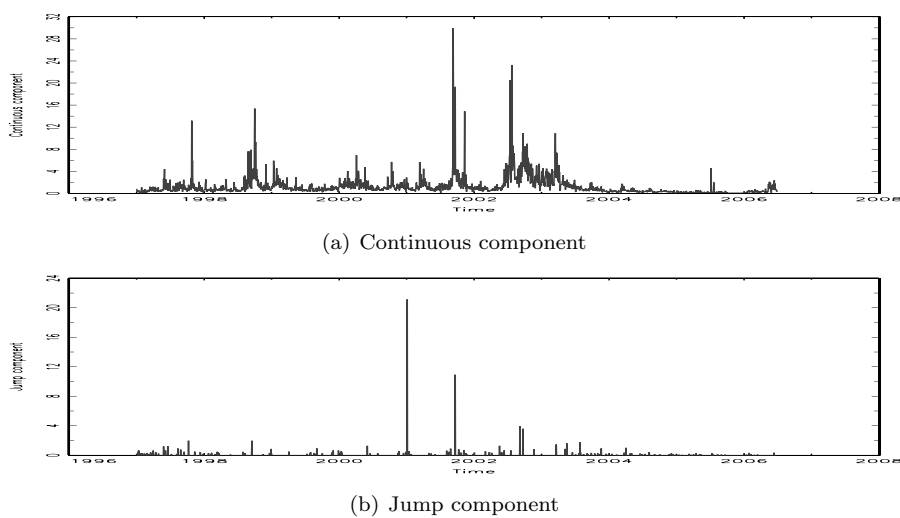
**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

Figure 2.5: Germany: Continuous and jump component



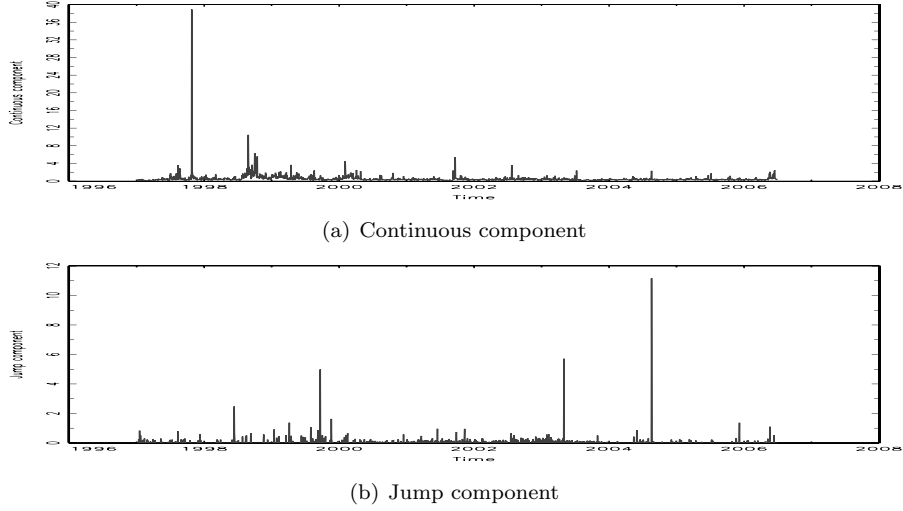
**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

Figure 2.6: France: Continuous and jump component



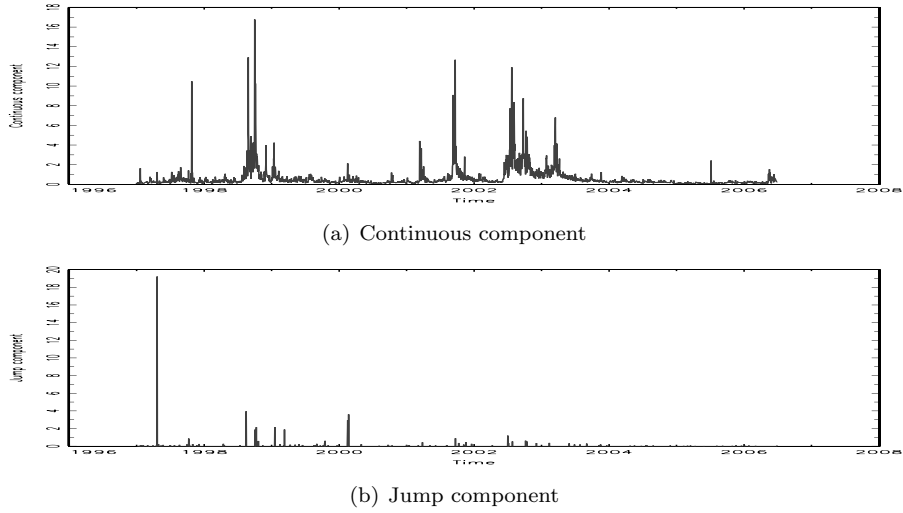
**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

Figure 2.7: Austria: Continuous and jump component



**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

Figure 2.8: Switzerland: Continuous and jump component



**Note:** The continuous component and jump component have been calculated according to the methods in Section 2.2. Missing values due to non-trading days or missing observations in the continuous component have been linearly interpolated.

test, the continuous component is never found to follow a unit root process. The autocorrelograms of all series (not included here), though, hint at long memory behavior and possibly fractional integration. For similar findings see, for example, Andersen et al. (2001).

Table 2.1: Descriptive statistics

	UK		US		NL		ITA	
	CC	JC	CC	JC	CC	JC	CC	JC
Mean	0.72	0.30	0.76	0.41	1.20	0.69	0.97	0.47
Median	0.42	0.17	0.51	0.21	0.58	0.27	0.55	0.21
Rel. no.	-	0.11	-	0.04	-	0.06	-	0.06
Maximum	31.32	11.18	11.28	5.41	6.00	31.93	35.59	7.66
Minimum	0.01	0.01	0.01	0.04	0.02	0.02	0.01	0.02
Std.Dev.	1.28	0.73	0.91	0.73	1.83	2.76	1.58	0.89
Skewness	10.77	13.22	4.97	4.61	4.42	10.52	8.56	5.24
Kurtosis	193.21	196.22	41.20	27.54	34.28	118.41	131.54	36.57
P-value (JB)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P-value (ADF)	0.00	-	0.00	-	0.00	-	0.00	-

	GER		FRA		AUT		CH	
	CC	JC	CC	JC	CC	JC	CC	JC
Mean	1.86	0.49	1.19	0.58	0.50	0.27	0.59	0.52
Median	1.05	0.29	0.74	0.25	0.36	0.15	0.30	0.11
Rel. no.	-	0.08	-	0.07	-	0.18	-	0.05
Maximum	51.38	9.11	29.95	21.13	38.90	11.17	16.77	19.22
Minimum	0.03	0.03	0.03	0.02	0.02	0.03	0.03	0.02
Std.Dev.	2.88	0.85	1.71	1.86	0.92	0.68	1.04	1.95
Skewness	6.78	7.14	6.50	9.19	30.08	11.87	6.59	8.53
Kurtosis	81.12	65.49	72.54	95.79	1209.06	170.74	66.38	81.45
P-value (JB)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P-value (ADF)	0.00	-	0.00	-	0.00	-	0.00	-

**Note:** Countries' names are abbreviated as: United Kingdom (UK), United States of America (US), the Netherlands (NL), Italy (ITA), Germany (GER), France (FRA), Austria (AUT), and Switzerland (CH). Rel.no. and Std.Dev. stand for relative number and standard deviation, respectively. The p-values of the Jarque-Bera (JB) and the augmented Dickey-Fuller (ADF) test are in %. We used a lag-length determination using the Schwarz information criterium.

### 2.3.2 Common factors in the volatility

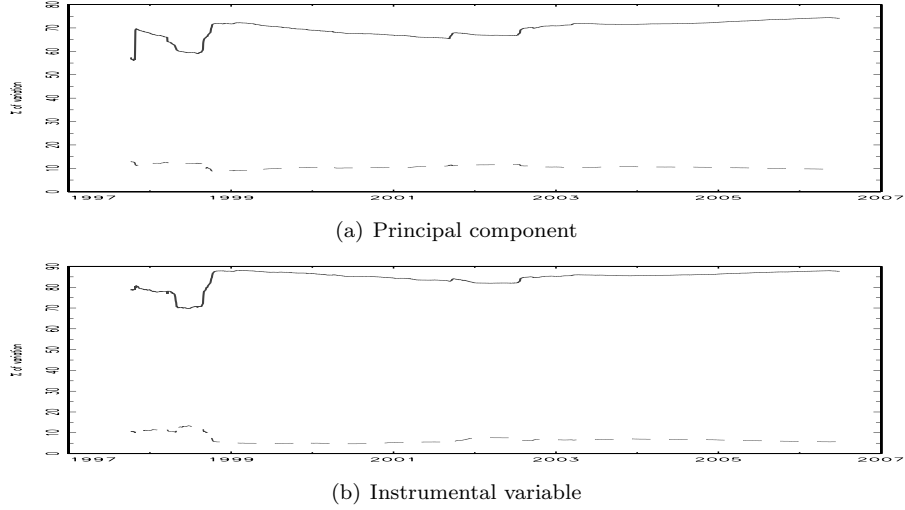
We already showed some basic results on the decomposition of the individual countries' stock market volatilities into a continuous and a jump component in Section 2.3.1 and Figures 2.1 to 2.8. There we notice that the individual countries' stock market volatilities exhibit some very pronounced periods of high volatility which also tend to coincide across markets. In the following two sections we would like to analyze these comovements of the two different parts of the assets' volatilities in more depth.

In Figures 2.9 to 2.11 we present the common factors representation of the

continuous volatility component as introduced in Anderson and Vahid (2007) and in Section 2.2 of this chapter. Such a factor representation can be justified by intuition but also by formal theories in finance which suggest that asset market (co)movements are driven by underlying market factors. For very good surveys on the theoretical and empirical literature on market factors driving asset movements see, for example, Chapter 6 in Campbell et al. (1997) and Chapter 9 in Cochrane (2001). Following Candelon et al. (2008a; 2008b) and because of likely differences in common factors, we split the sample into three different subsets in order to see if there are any discrepancies across them. The first set is the full sample of countries consisting of the UK, US, France, Germany, Italy, the Netherlands, Austria and Switzerland. A second subset is a “pure” European sample without the US. In a last subset we also exclude Austria and Switzerland giving us a sample of European countries which we call European core countries consisting of the UK, France, Germany, Italy, and the Netherlands. All the three figures show the first two identified factors and their relative contribution to explaining the overall variation in the sample. Panel A represents the principal component and Panel B the instrumental variable approach as suggested in Anderson and Vahid (2007). As explained above a factor is deemed significant if it is able to explain at least as much as an “average” factor would do. An “average” factor would be a factor that explains a fraction of the total variation in the data equal to  $\frac{100\%}{\text{number of possible factors}}$ , where the number of possible factors is equal to the number of markets or series in the considered sample. We prefer this rather heuristic approach over the factor selection criteria proposed by Bai and Ng (2002) because of our relatively small  $N$  (amount of series in the data set). Bai and Ng (2002), Engle and Marcucci (2006), and Anderson and Vahid (2007) find that the model selection criteria select a large number of common factors relative to  $N$  when  $N$  is small. We nevertheless calculated the criteria proposed by Bai and Ng (2002) and found that also in our case they would almost always have selected the maximum amount of factors, which economically does not make much sense.

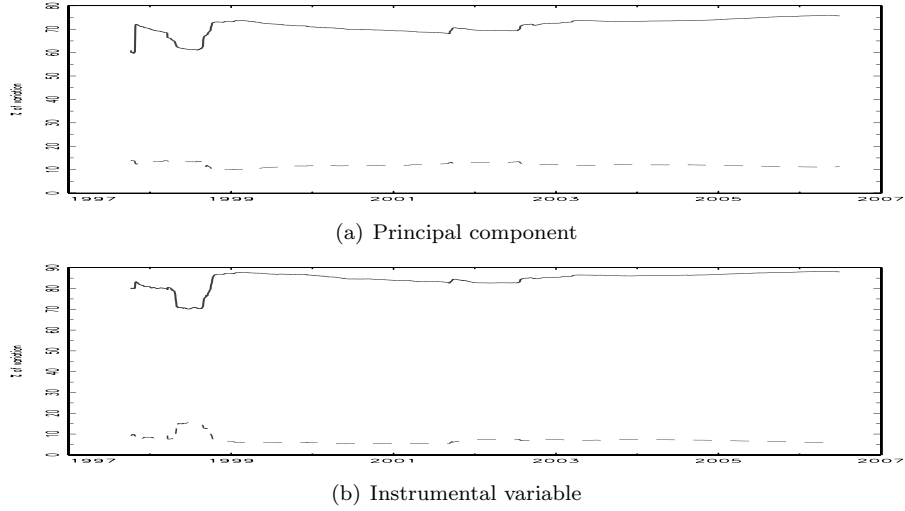
According to our definition of significance of factors we find that in all considered cases only one factor is found to be decisive although a second one often is relatively close to being significant as well. Therefore, we show in Figures 2.9 to 2.11 both the first and the second most important factor, because we find it interesting to also see the difference in importance of the factors graphically. In the analysis we further focus only on the first common factor. In general, one can see in all three subsamples and figures that there is an increase in the importance of the first common factor from the beginning to the end of the sample period. In the first two years of the sample around 60 – 70% and at the end of the sample around 75 – 85% of the sample variation is accounted for by the first common fac-

Figure 2.9: Common factors CC: Europe core, US, Austria and Switzerland



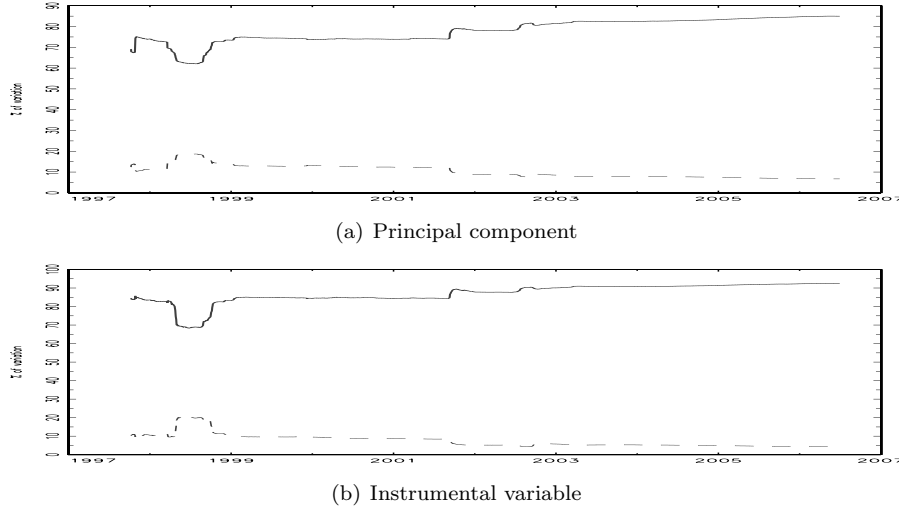
**Note:** The common factors have been calculated according to the methods in Section 2.2. The solid line represents the first, the dashed line represents the second common factor.

Figure 2.10: Common factors CC: Europe core, Austria and Switzerland



**Note:** The common factors have been calculated according to the methods in Section 2.2. The solid line represents the first, the dashed line represents the second common factor.

Figure 2.11: Common factors CC: Europe core



**Note:** The common factors have been calculated according to the methods in Section 2.2. The solid line represents the first, the dashed line represents the second common factor.

tor. Despite this “upward-trend” in the common factor there are some interesting observations to be made.

For all subsets of countries in the beginning of the sample period there seems to be quite strong fluctuation in the importance of the common factors. Compared to the principal component calculations, the first factor based on the IV estimation starts off relatively high. Then in the beginning of 1998 it declines steadily to increase steeply again at the end of 1998. Such a behavior can be expected by looking at Figures 2.1 to 2.8 and observing the clear increases in the continuous component of the volatility at the end of 1997 and 1998 which is shared by virtually all markets. An economic explanation for these spikes in volatility is the occurrence of the East-Asian crises in the second half of 1997 and the Russian crisis that hit in August 1998 which was exacerbated by a global recession in 1998. Especially European countries were struck by the Russian crisis which led to a prolonged period of very volatile stock markets in Europe. A word of caution is appropriate here. In the beginning of the sample period stronger fluctuations of the common factor importance are to be expected because the length of the data set at that point is short by construction. Gradually extending it as the sample progresses will on average lead to a higher variance of the common factor than for the full sample. Nevertheless, we believe that the economic events given above are the

main cause for the movements in the factor, which is supported by Figures 2.1 to 2.8.

In Figure 2.9 and 2.10 one can see that after the increase due to the Russian crisis, the importance of the factor starts to gradually decline until stepping up again in September 2001 and the second half of 2002. Later the importance further increases gradually until the end of the sampling period. The two steps in the first two years of the new millennium can also easily be spotted in Figures 2.1 to 2.8 with the exemption of Austria. In September 2001 the terrorist attacks on the World Trade Center in New York led to worldwide uncertainty reflected by increased asset market volatilities. Also in the second half of 2002 and the beginning of 2003 the world experienced increased political uncertainty leading to elevated volatility. This period was affected by the build-up to the second Gulf War against Iraq. Virtually all countries were affected by these events leading to the described increases in the importance of the common factor. After this, the further gradual increase in the relative variance explained by the first common factor is most likely to be caused by the historically (by sample period standards) low volatilities in the stock markets across all markets in the sample.

Only considering the European core countries, an even clearer upward-trend in the importance of the first common factor for both the principal component and the instrumental variable estimators can be seen. Also the level of total variance explained by this factor is shifted upward compared to the samples including the US, Austria and Switzerland. The upward movements of the factor's importance is also not interrupted by a downward movement between 1998 and 2001 as in the larger subsamples. So, with the core European countries there is a clear trend towards increased importance of the first common factor only interrupted by strong non-gradual increases in 1998, 2001, and 2002 comparable to the cases in Figures 2.9 and 2.10. Another interesting point to note for the European core countries is that there does not seem to be any significant reaction of the common factors to the introduction of the euro.

In sum, one can say that only one common factor is significant in all considered cases. This factor's importance increases on average during the sample period with some periods characterized by "jump" behavior, especially during international financial crises and global economic uncertainty in 1997, 1998, 2001 and 2002/2003. Particularly in the core European countries the increase in the importance of the common factor during the sample period is evident. The behavior of the common factor found in the data is likely to be explained by major international economic and political events at least at the indicated points in time. The apparent upward-trend in the importance of the common factor cannot definitely be explained here. An intuitive candidate, especially for the core European countries, is the continu-



ation of the (European) economic integration process and the implementation of the common market. Such a discussion will be shortly revisited in Chapter 4 of this thesis where we analyze possible structural breaks in the concordance of bulls and bears markets in East-Asia.

### 2.3.3 Co-jumps

In Section 2.3.2 we analyzed the continuous component of the volatility process and its international linkages. Here we would like to focus on the jump component of the process. In Section 2.2 and 2.3.1 respectively, we showed theoretically and empirically the separation of the volatility of stock market price processes into their continuous and their jump component. Depending on a chosen jump significance level  $\alpha$  we identify days on which there are significant jumps in the price process. Obviously, we can then proceed to identify the exact high-frequency returns during the trading day which cause the jump test statistic to be statistically significant. By doing so we follow a very intuitive approach that is explained in Section 2.2 and also used in Beine et al. (2007).

While in Section 2.3.1 we already had a look at the jump component for every individual country in the sample, we want to extend that analysis here to possible co-jumps across countries. As explained in Section 2.2 we define co-jumps as either positive or negative co-jumps when two countries show either a positive or negative jump on the same day, respectively.<sup>15</sup> The results of such a co-jump analysis are summarized in Table 2.2, where we show the positive and negative co-jump sample probabilities for every country combination. For example, the 0.77% in Panel A in the column for the Netherlands and the row of the UK means that there is a 0.77% chance every day, or once every 130 trading days, that we observe a positive jump in both countries on that day. In Panel A and B we use a significance level for the jump test statistic of  $\alpha = 99.9\%$  and  $\alpha = 99.99\%$ , respectively.

In the case of independent international stock markets one would expect that co-jumps across those markets are the exception rather than the rule. However, once one acknowledges that international stock markets are interconnected, such co-jumps become a reasonable possibility. A dependence of international stock markets in terms of their jump behavior has implications for traders trying to hedge their stock market risks by international diversification. Such a diversification strategy would be less effective if countries' markets tend to co-jump. Also financial regulators are interested in such phenomena in case they would like to regulate financial markets more heavily to prevent excessive comovements. The area of

---

<sup>15</sup>We also had a look at cases where one country shows a positive and another country a negative jump. These cases are very few and we therefore neglect them here. They are nevertheless available from the authors upon request.

high-frequency co-jumps is a very new area of research with not many contributions yet. Notable exceptions are, for example, Gobbi and Mancini (2006), Lahaye et al. (2007) and the references therein.

Table 2.2: Co-jumps probability in %

Panel A:  $\alpha = 99.9\%$ 

		Positive Co-Jumps							
Neg. Co-Jumps		UK	US	NL	ITA	GER	FRA	AUT	CH
	UK		0.48**	0.77**	0.61*	0.48 <sub>a</sub>	0.73 <sub>a</sub>	0.44*	0.12
	US	0.16		0.16 <sub>b</sub>	0.24 <sub>a</sub>	0.24 <sub>a</sub>	0.32*	0.16	0.00 <sub>a</sub>
	NL	0.44 <sub>a</sub>	0.12		0.44 <sub>a</sub>	0.68 <sub>a</sub>	0.73*	0.48*	0.20 <sub>a</sub>
	ITA	0.36 <sub>a</sub>	0.16 <sub>b</sub>	0.40 <sub>a</sub>		0.48 <sub>a</sub>	0.65*	0.24	0.12
	GER	0.61 <sub>a</sub>	0.20 <sub>b</sub>	0.73 <sub>a</sub>	0.44 <sub>a</sub>		0.77 <sub>a</sub>	0.48	0.16
	FRA	0.73 <sub>a</sub>	0.16 <sub>b</sub>	0.52 <sub>a</sub>	0.48 <sub>a</sub>	0.65 <sub>a</sub>		0.48 <sub>b</sub>	0.12
	AUT	0.24	0.12	0.28	0.24	0.48	0.44 <sub>b</sub>		0.32*
	CH	0.28**	0.12*	0.20 <sub>a</sub>	0.20 <sub>a</sub>	0.40**	0.24*	0.20	

Panel B:  $\alpha = 99.99\%$ 

		Positive Co-Jumps							
Neg. Co-Jumps		UK	US	NL	ITA	GER	FRA	AUT	CH
	UK		0.16**	0.40*	0.28**	0.28 <sub>a</sub>	0.28 <sub>a</sub>	0.16*	0.00 <sub>a</sub>
	US	0.00 <sub>a</sub>		0.08 <sub>b</sub>	0.12 <sub>a</sub>	0.04	0.16**	0.00 <sub>a</sub>	0.00 <sub>a</sub>
	NL	0.24 <sub>a</sub>	0.00 <sub>a</sub>		0.36**	0.36*	0.40**	0.12*	0.08 <sub>a</sub>
	ITA	0.08	0.12 <sub>a</sub>	0.08 <sub>a</sub>		0.24 <sub>a</sub>	0.28*	0.08	0.04
	GER	0.32 <sub>a</sub>	0.12*	0.20 <sub>a</sub>	0.16 <sub>a</sub>		0.32 <sub>a</sub>	0.24*	0.08 <sub>b</sub>
	FRA	0.36 <sub>a</sub>	0.00 <sub>a</sub>	0.20 <sub>a</sub>	0.12 <sub>a</sub>	0.32 <sub>a</sub>		0.16	0.08 <sub>b</sub>
	AUT	0.08	0.00 <sub>a</sub>	0.04	0.08	0.12	0.20 <sub>b</sub>		0.04
	CH	0.12**	0.04	0.08 <sub>a</sub>	0.04	0.12 <sub>a</sub>	0.04	0.08	

**Note:** Countries' names are abbreviated as: United Kingdom (UK), United States of America (US), the Netherlands (NL), Italy (ITA), Germany (GER), France (FRA), Austria (AUT), and Switzerland (CH). For the symmetry null hypothesis \* means statistical significance at a 10%, \*\* at a 5% level. For the independence null hypothesis *b* means statistical significance at a 10%, *a* at a 5% level.

Therefore, we test if the event of co-jumps is dependent or independent on the information that at least one of the two countries shows a jump. Dependence here means that the knowledge that one country has a positive (negative) jump changes the probability that the other country also shows a positive (negative) jump compared to its unconditional probability. So, rejecting the null hypothesis of independence can be taken as evidence that there is a linkage between those markets. In order to test the null hypothesis of independence of the jump occurrences in both countries we have to assume a distribution for the jump occurrence in each country. We assume that jumps follow a binomial distribution. Either a day shows a jump (success) or it does not (no success).<sup>16</sup> As estimates for the probability of showing a jump we take  $p = \hat{p}_i^{+, -}$ , where *i* stands for country *i* and

<sup>16</sup>We assume two different binomial distributions for positive and negative jumps for each country.

+ or – stand for positive and negative jumps, respectively. As estimates we take the observed sample frequencies. Under the null hypothesis of independence we have:

$$p(Z_{i,t}^{+,-} > \phi(\alpha), Z_{j,t}^{+,-} > \phi(\alpha)) = p(Z_{i,t}^{+,-} > \phi(\alpha))p(Z_{j,t}^{+,-} > \phi(\alpha)), \quad (2.15)$$

where  $\phi(\alpha)$  stands for the abscissa value of the standard normal distribution at the significance level  $\alpha$ . In other words, under the null the daily probability of having a co-jump being either positive (+) or negative (-) should be equal to the product of the two marginal distributions of having such a jump in country  $i$  and  $j$ , respectively. With these assumptions we can test the null hypothesis of independence. Results of statistical significance are reported in Table 2.2 as subscript  $a$  or  $b$  corresponding to a 5% and 10% significance level, respectively.

In almost all cases the null hypothesis of independence is rejected in favor of the alternative that the probability of having a co-jump is significantly larger than it would be under independence. Only the cases involving Austria and Switzerland show some different results, which are also driven by the fact of some zero-observations especially in Panel B of Table 2.2. Nevertheless, the general picture is that the information of country  $i$  showing a positive or negative jump significantly increases the chance of also observing a positive or negative jump on the same day in country  $j$ .

Considering the literature on asymmetries in marked movements, like Bekaert and Wu (2000), Bollerslev et al. (2006) and the references therein, one might wonder if there are also asymmetries in the international stock market co-jump behavior. We therefore ask the question if stock market co-jumps are more likely to occur when they are negative than when they are positive. Finding asymmetric co-extreme return behavior would mean that there is larger downward risk than upward potential in an internationally diversified portfolio once we focus on the jump component.

In Table 2.2 it is apparent that most values in the upper triangular are larger than in the lower triangular. In order to test for a statistically significant difference between both we apply a similar procedure as above in the test for independence. We highlight that country pair, either the probability of positive or the probability of negative co-jumps, which is statistically more likely to occur. We clearly find that most of the cases where we find significant differences in the probabilities, the positive co-jumps are more likely than the negative co-jumps.<sup>17</sup> By only looking at the jump component it thereby appears as if the upward potential in an internationally diversified portfolio is larger than the downward risk.

---

<sup>17</sup>Again we see that Switzerland stands out a bit as a special case.

### 2.3.4 Close-open returns and jumps

Up to now we have focused on the intraday trading hour returns. In order to round up the analysis we also need to have a look at the “interday” or close-open returns and their relationship with the intraday returns. Andersen et al. (2007) already found that the close-open or overnight returns constitute an important part of the total variation of returns as most financial assets’ prices change over night. They also find that the dynamics of the overnight return volatility differs from those of the continuous and the jump component.<sup>18</sup> As stock markets are usually closed over night we do not have any high-frequency data for the periods from market closing until reopening on the following trading day. Giot et al. (2007) state that for example low trading volume increases the likelihood of observing jumps. In our work we are rather interested in the overnight returns themselves and their possible relation with jumps on the following trading day (this subsection) and across international stock markets (next subsection). Mostly investors do not buy and sell their assets on the same trading day but hold them over longer horizons. Also do many investors hold assets in different countries in order to diversify their portfolios. Such investors’ portfolio returns and risks will be influenced by (extreme) overnight returns, their international interconnection and their relationship with asset price jumps.

According to our best knowledge there is no study yet which combines these two measures. Here we are primarily interested if the probability of observing one or more jumps during trading hours is related to the size of the close-open return realized at the beginning of that day. One might expect that news causing extreme close-open returns tends to also cause an increased intraday jump behavior because investors are struggling to determine the exact impact of the news on stock prices. Such an uncertainty would easily transform into more pronounced jump behavior after “big news”.

Usually, new information, arriving when stock markets are closed, leads to differences between the closing time price on day  $t - 1$  and the opening price on day  $t$ . So, depending on the importance of the news arriving during the non-trading period, opening prices will differ from closing prices leading to differences in the so-called close-open returns. Under the assumption of independence of close-open returns and the probability of having a jump in the intradaily price process on that day, there should not be a difference between the overall likelihood of having a jump on any day and the likelihood of finding a jump following extremely positive (negative) close-open returns. Furthermore, an efficient market hypothesis as in Fama (1970) and Malkiel (1987) states that jumps should not be predictable,

---

<sup>18</sup>Andersen et al. (2007) are, among others, interested in the overnight returns and their contribution to the total return variability.

because otherwise there were arbitrage opportunities. So, by testing the hypothesis of independence between overnight returns and the propensity for jumps on the following trading day, we can indirectly test if the efficient market hypothesis holds. It is therefore interesting to know if the size of the close-open return realized directly at the beginning of each trading day has any predictive power for the likelihood of increased jump-behavior during the following trading hours. Such information is important for high-frequency trading and also for option pricing where the continuous component and the jump component of the volatility can have different impacts.<sup>19</sup>

Table 2.3: Extreme close-open returns and jumps

Panel A:  $\alpha = 99.9\%$ 

	Full Sample			Without Mondays		
	All obs.	LP	UP	All obs.	LP	UP
UK	0.111	0.054**	0.077	0.107	0.039**	0.077
US	0.041	0.000**	0.033	0.042	0.000**	0.021
NL	0.058	0.043	0.026	0.060	0.022*	0.033
ITA	0.061	0.044	0.044	0.063	0.034	0.045
GER	0.079	0.041*	0.008**	0.080	0.042	0.010**
FRA	0.069	0.058	0.033**	0.069	0.064	0.042
AUT	0.164	0.203	0.118	0.162	0.187	0.13
CH	0.042	0.033	0.083	0.038	0.000**	0.053

Panel B:  $\alpha = 99.99\%$ 

	Full Sample			Without Mondays		
	All obs.	LP	UP	All obs.	LP	UP
UK	0.061	0.041	0.049	0.059	0.031	0.041
US	0.020	0.000*	0.008	0.022	0.000	0.000
NL	0.028	0.009	0.009	0.031	0.011	0.022
ITA	0.029	0.018	0.026	0.032	0.011	0.022
GER	0.033	0.025	0.000**	0.035	0.032	0.010
FRA	0.036	0.025	0.017	0.036	0.022	0.032
AUT	0.100	0.144	0.059*	0.095	0.132	0.065
CH	0.022	0.008	0.041	0.020	0.000	0.042

**Note:** Countries' names are abbreviated as: United Kingdom (UK), United States of America (US), the Netherlands (NL), Italy (ITA), Germany (GER), France (FRA), Austria (AUT), and Switzerland (CH). All obs., LP, and UP stand for all observations, lower percentile, and upper percentile, respectively. \* means statistical significance at a 10%, \*\* at a 5% level.

In Table 2.3 we summarize the results of such an exercise. We distinguish here between the “full sample” and the case without Mondays and trading days after days without trading. Such a distinction is important because there are potentially large differences in the amount of price-sensitive news arriving during the night and during weekends and holidays. As references about weekend effects see, for example, Cross (1973), Rogalski (1984), and Abraham and Ikenberry (1994).

<sup>19</sup>See, for example, Stentoft (2008).

So, the close-open returns on Mondays incorporate news from the closing on the Friday before until the opening on Monday, whereas the usual close-open return on the other days of the week only correspond to news accumulated within say 16 hours. Accounting or not accounting for such effects obviously changes the unconditional distribution of close-open returns and potentially their tail behavior. Another distinction considered in the table is between large negative (lower percentile or LP) and positive (upper percentile or UP) close-open returns. Our definition of an “extreme” observation is that the close-open return has to be in the lower (LP) or upper (UP) 5% percentile of the unconditional distribution of close-open returns.<sup>20</sup> Obviously, there might be a difference in the markets’ reactions to extremely good or extremely bad news. The column “All observations” summarizes the unconditional probability of observing a jump on any trading day. Lastly, the table separates the results between two different significance levels for the jump test statistic in Equation 2.9. The upper panel uses  $\alpha = 99.9\%$ , the lower one uses  $\alpha = 99.99\%$  as significance levels, obviously reducing the amount of jumps found during trading hours.

Under the null hypothesis of independence, the figures in the columns “all observations” “LP” and “UP” should be the same up to sampling variation for every individual market and considered sample. In general, Table 2.3 shows that all countries but Austria and Switzerland have lower probabilities of observing a jump during a trading day given that the foregoing close-open return was either extremely negative or positive.

In order to test for the statistical significance of those differences in jump probabilities we have to make some distributional assumptions under the null hypothesis of independence between the size of close-open returns and the likelihood of jumps on the following trading day. It is reasonable to assume that the occurrence of a jump on any given trading day is distributed according to the binomial distribution with parameters  $N$  being the sample size and  $p$  being the “success” probability of observing a jump on any given trading day. Such a binomial distribution is reasonable because the jump statistic is also constructed under the assumption of independence of past price observations. If future jumps in the price process were predictable on the basis of a given information set known to the market, arbitrage opportunities would exist and would immediately lead to corresponding price adjustments. As an estimate for the probability  $p$  we take the unconditional probability for observing a jump using all observations in the sample. With these assumptions at hand we can easily check if the sample-frequencies of jumps given

<sup>20</sup>Taking a cut-off value even more extreme, say 1% or even smaller, would already lead to regions where one would probably have to use Extreme-Value-Theory (EVT). Such an approach would not add much to the findings at this point. In Chapter 5 of this thesis we perform such an analysis in a somewhat different context though.

either an extremely negative or positive close-open return differ significantly from those using the whole distribution of close-open returns. We indicate significance by adding stars to the entries in the table. One star indicates significance at a 10%, two stars significance at a 5% level.

For the case with  $\alpha = 99.9\%$  and using the full sample, so including Mondays and trading days after holidays, we find five of the possible 16 cases to be significant. Those are the lower tails of the UK, US and Germany, and the upper tails of Germany and France. The results for the reduced sample excluding Mondays and trading days after holidays are similar. So, there is only minor evidence that trading days after weekends and holidays show different behavior than “normal” trading days.

In the case of  $\alpha = 99.99\%$  we generally find less significant cases in the table. For the full sample only the lower case of the US and the upper percentile case of Germany remain significant. Also the upper percentile for Austria turned significant here, but only at a 10% level. For the reduced sample no cases are significant anymore. Such an outcome is not surprising because increasing the jump significance level  $\alpha$  obviously reduces the amount of identified jumps in the first place reducing the probability of observing jumps on any given trading day. At a certain point even observing zero jumps on days corresponding to the upper or lower percentile of close-open returns would not be found to be a statistically rare event anymore.

In sum, we can say that there is some significant evidence that intraday jumps are less frequent after extreme close-open returns than if we did not condition on the size of the close-open return. Such a finding at least casts some doubts on an efficient market hypothesis saying that future asset returns are not predictable using current information.

### 2.3.5 Close-open returns across countries

Having analyzed the close-open returns in a within-country setting we now want to focus on those returns across markets. We are again interested in the upper and lower 5% percentiles of the close-open returns distributions of the individual countries. But here we want to analyze if those extreme negative and positive observations tend to occur simultaneously across different countries’ stock markets. In other words, what is the probability of, for example, two countries having an extreme increase in stock prices during non-trading hours given that one of the two countries does so. We also test if negative co-extremes are more likely to occur than positive ones or vice versa, which would mean that there are asymmetries in the co-extreme close-open return behavior across countries.

Industrialized countries’ stock markets are very much integrated with each

other.<sup>21</sup> As such we expect strong co-extreme behavior in the close-open returns, because in integrated markets extreme shocks are expected to be transferred among each other. In the case of asymmetric behavior our prior is that negative co-extremes are more likely than positive ones, because usually investors are more responsive to extreme negative news than to positive ones. Asymmetries in stock markets' price processes have, for example, been dealt with in Bekaert and Wu (2000), Bollerslev et al. (2006) and the references therein. Possible asymmetries can be important for short term investors who are diversified into different asset markets. Finding asymmetric co-extreme return behavior would mean that there is larger downward risk than upward potential in such a portfolio considering extreme close-open returns.

Table 2.4: Co-extremes in close-open returns (unadjusted)

Panel A: Full sample

		Upper percentiles							
Lower percentiles	UK	US	NL	ITA	GER	FRA	AUT	CH	
	UK		0.140	0.559	0.352	0.520	0.540	0.207	0.606
	US	0.212**		0.125	0.140	0.176	0.146	0.047	0.133
	NL	0.522	0.230**		0.447	0.598	0.629	0.167	0.637
	ITA	0.368	0.188	0.536		0.382	0.494	0.184	0.398
	GER	0.535	0.233	0.571	0.489**		0.527	0.170	0.526
	FRA	0.606*	0.216	0.682	0.534	0.598*		0.154	0.553
	AUT	0.220	0.082*	0.236*	0.256	0.301**	0.278**		0.174
	CH	0.602	0.214	0.678	0.517**	0.585	0.591	0.242*	

Panel B: Without Mondays and days following holidays

		Upper percentiles						
Lower percentiles	UK	US	NL	ITA	GER	FRA	AUT	CH
	UK	0.191	0.530	0.342	0.519	0.516	0.145	0.590
	US	0.224	0.212	0.216	0.256	0.235	0.088	0.239
	NL	0.569	0.304**	0.426	0.603	0.602	0.195	0.635
	ITA	0.373	0.246	0.533*	0.434	0.509	0.207	0.420
	GER	0.531	0.285	0.617	0.455	0.509	0.177	0.542
	FRA	0.568	0.237	0.696**	0.532	0.598**	0.181	0.546
	AUT	0.224**	0.062	0.250	0.246	0.263**	0.261**	0.203
CH	0.553	0.241	0.684	0.532**	0.588	0.610	0.239	

**Note:** Countries' names are abbreviated as: United Kingdom (UK), United States of America (US), the Netherlands (NL), Italy (ITA), Germany (GER), France (FRA), Austria (AUT), and Switzerland (CH). \* means statistical significance at a 10%, \*\* at a 5% level.

We summarize our results in Tables 2.4 and 2.5. In order to account for different trading hours we present two tables here. Table 2.4 shows the results for the “unadjusted” close-open returns, which means that we calculate the returns as they come for every individual country and then calculate the probabilities of co-occurrences. These returns do not take the different opening and closing hours

<sup>21</sup>See, for example, Kim et al. (2005) and Candelon et al. (2008a).



across countries into account. A potentially important factor, though, is that if one country opens before the second one, the opening price of the second country already incorporates information from the opening and trading in country one (the same for the closing time). In order to adjust for this we also report in Table 2.5 results where we standardize the closing and the opening times for every country pair. Here we take as closing price those observations with the latest available time stamp in both markets. For the opening price observations we use the first time stamp that is available in both markets. Herewith we make use of the advantage of having high frequency intra-day data which enables us to “adjust” the close-open returns in such a way that they are not distorted by different trading hours.<sup>22</sup>

Table 2.5: Co-extremes in close-open returns (adjusted)

Panel A: Full sample

		Upper percentiles						
Lower percentiles	UK	US	NL	ITA	GER	FRA	AUT	CH
	UK	0.460	0.559	0.421	0.480	0.520	0.272	0.495
	US	0.414	0.4713	0.349	0.409	0.393	0.200	0.421
	NL	0.554	0.454	0.388	0.620	0.640	0.322	0.637
	ITA	0.517*	0.365	0.602	0.461	0.551	0.379	0.409
	GER	0.555*	0.460	0.681	0.648**	0.656	0.160	0.621
	FRA	0.626**	0.466	0.693	0.580	0.717	0.275	0.575
	AUT	0.341	0.271*	0.337	0.442	0.269**	0.422**	0.207
	CH	0.553	0.414	0.733**	0.506*	0.649	0.613	0.286*

Panel B: Without Mondays and days following holidays

		Upper percentiles						
Lower percentiles	UK	US	NL	ITA	GER	FRA	AUT	CH
	US	0.440	0.446	0.460	0.421	0.426	0.203	0.407
	NL	0.595**	0.487*	0.407	0.621	0.611	0.336	0.635
	ITA	0.509	0.330	0.579**	0.496	0.598	0.405	0.464
	GER	0.523	0.460	0.687	0.589**	0.644	0.168	0.600
	FRA	0.576	0.465	0.696*	0.568	0.735**	0.310	0.563
	AUT	0.345**	0.256	0.339	0.482*	0.263**	0.417**	0.263
	CH	0.580**	0.411	0.702	0.514	0.622	0.653**	0.308

**Note:** Countries' names are abbreviated as: United Kingdom (UK), United States of America (US), the Netherlands (NL), Italy (ITA), Germany (GER), France (FRA), Austria (AUT), and Switzerland (CH). \* means statistical significance at a 10%, \*\* at a 5% level.

The two tables report the frequency of “co-extremes” either in the lower or the upper percentile of the countries’ close-open return distributions relative to the maximum amount of possible co-events. For example, in Table 2.4 Panel A the entry in the column “UK” and the row “FRA” is equal to 0.606. This means that in the sample in 60.6% of the possible cases the UK and France had an extreme negative close-open return (lower 5% percentile of their respective unconditional

<sup>22</sup>Trading hours in Europe do not differ very much. Exact closing and opening times are available from the authors upon request.

distributions) on the same day. If the events of observing a close-open return in the UK and France in their lower 5% percentiles were independent, such a probability should average 2.5‰ instead of 60.6%. As expected, all reported figures are significantly larger than 2.5‰ and independence can thereby be rejected. This confirms the well-known results in the literature that asset-markets tend to be positively correlated especially within the same economic region.<sup>23</sup>

As expected by the asymmetry hypothesis, the figures in the lower triangular are almost always larger than the corresponding figures in the upper triangular. This means that on average the probability of observing negative co-extremes is higher than the probability of observing positive co-extremes. In order to check for statistical significance of these asymmetries we performed a bootstrap simulation on the close-open returns. Stars indicate that there is a significant difference in the two co-exceedances. On average, in ca. 40% of the cases the negative co-exceedance probability is significantly larger than those in the upper percentiles. In none of the cases it is the other way around. Austria again shows different results with on average significantly smaller cross-country probabilities for both the upper and the lower percentiles than all the other pairs of countries.<sup>24</sup>

Such strong and apparently mostly asymmetric behavior in the comovement of European and US stock markets in the sample period can be a sign of strong international linkages of the markets through, for example, common macroeconomic fundamentals. Other possible explanations are based on investors' behavior that can also be asymmetric, meaning that bad news tends to cause stronger movements out of stocks on an international scale than good news tending towards movements into stocks. A topic that we will partly come back to in Chapter 5 of this thesis. Such results are especially interesting in the light of the findings for the co-jumps where we found that positive co-jumps are more likely to occur than negative ones. In general, it is difficult to draw conclusions at this stage if these (asymmetric) comovements are caused by strong inter-country linkages being macroeconomic or financial, or if they are caused by global factors affecting all countries at the same time. Here we rather want to deliver some stylized facts to highlight the importance of the comovement behavior of financial markets than analyzing possible causes. We leave such an analysis for future research.

<sup>23</sup>See, for example, Candelon et al. (2008a) and the references therein.

<sup>24</sup>Another notable exception is the US in the unadjusted case. Comparison with the US results in the adjusted returns case, though, shows that this is probably due to the fact that the overlap of the US trading hours with those of the other European countries is relatively short and thereby potentially leads to large differences between adjusted and unadjusted close-open returns.

## 2.4 Conclusions

In this chapter we gave a brief introduction to the high-frequency analysis of stock markets. We first focused on disentangling the continuous and the jump components from price diffusion processes with the help of five minute high-frequency stock index observations for a sample of eight mostly European industrialized countries. We paid attention to every country in isolation in order to extend the analysis towards finding linkages among the countries' stock markets with respect to close-open returns, the continuous component of volatility, jumps and co-jump behavior.

We find that the extracted series for the continuous volatility component and the jump component differ a lot across countries especially in terms of their unconditional distributions. Nevertheless, we see that much of the variation of the continuous component varies together across countries. This is then further supported by the finding that one significant common factor is able to explain up to ca. 85% of the overall sample variation in the continuous component of the volatility. Such a common factor is also found to trend upwards in its importance of explaining the sample variation. On the one hand major international economic, financial, and political events and crises are likely to have caused upward-shifts in the importance of the common factor. On the other hand we also see some gradual increase over most of the sample period especially for the core European countries UK, France, Germany, Italy, and the Netherlands. Such a gradual increase can be explained by an increasing importance of financial and economic integration among these countries.

We further find important and significant asymmetries with respect to the co-extreme behavior of close-open returns across countries. Negative co-extreme returns are found to be on average more likely to occur than positive ones. So, markets are more reactive to bad news during non-trading hours in a coordinated sense than with respect to good news. Such an asymmetry is also found in the case of co-jumps during opening hours of stock exchanges, but here the asymmetry is in the opposite direction. Positive co-jumps are significantly more likely to occur than negative co-jumps, which might be some sort of backlash for excessive negative close-open returns in the beginning of the trading day. At this stage it is difficult to draw any clear-cut conclusions about possible causes, though.

Countries' stock markets are clearly found to co-move in many aspects. We focused here on the first and second moment or the returns and volatilities of the return processes. Clearly other aspects of the (un)conditional return distributions might be interconnected or tend to co-move as well. Therefore, in the following chapters we will partly go into more detail of the co-moving behavior of asset markets across the world.

---

# MULTIVARIATE STOCK MARKET SYNCHRONIZATION IN SMALL SAMPLES: THE CASE OF EAST-ASIA

Here we implement estimation and testing procedures for comovements of stock market “cycles” or “phases” in Asia.<sup>1</sup> We extend the Generalized Method of Moments (GMM) approach to measuring business cycle synchronization due to Harding and Pagan (2006b) toward estimating and testing for (bivariate, multivariate) cyclical stock market synchronization in a small sample framework. We show that the asymptotic version of the test gets increasingly distorted in small samples when the number of countries grows large. However, a block bootstrap version of the test can remedy the size distortion when the time series length divided by the number of countries  $T/n$  is sufficiently large. Applying the financial cycle synchronization framework to five East-Asian economies, we are unable to reject the null hypothesis of a nonzero common multivariate synchronization index for certain economically meaningful subsets of these countries.

## 3.1 Introduction

In the past couple of years, financial economists have extensively documented the empirical features of stock market returns such as clusters of volatility and heavy tails, see e.g. Embrechts et al. (1997). That stock prices typically exhibit periods of persistent rises or falls, i.e., so-called “bulls” and “bears”, has been recognized by the financial practitioners for a long time but has attracted much less attention from the academic community.<sup>2</sup> Accordingly, the potential for stock markets to

---

<sup>1</sup>This chapter is based on the paper Candelon, Piplack, and Straetmans (2008a) and my work with Stephan Smeekes on the applicability of the blockbootstrap procedure.

<sup>2</sup>Traditionally, there are two perspectives on stock market bulls and bears. First, they may be induced by irrational “animal spirit” (or sentiment that is unrelated to any rational expectations

be simultaneously bullish or bearish across geographical borders has also stayed underexposed. The main objective of this chapter will therefore be to provide a framework for measuring synchronization between stock market cycles.<sup>3</sup>

Measuring stock market cycles and their cross-border synchronization is of potential interest for both investors and policy makers. First, it is common sense that investors rebalance their portfolios by purchasing “cheap” stocks during bearish periods and selling “expensive” stocks when stock markets are bullish. The question arises, however, how to optimally time this portfolio rebalancing. Technical analysts typically time their buying and selling decisions by means of graphs. A more thorough statistical analysis of bulls and bears can help investors to further improve the timing of their investment decisions. The above discussion suggests that the duration of a stock cycle constitutes the natural time horizon for a “single-cycle” or “short-term” investor (or, alternatively, constitutes the natural time horizon for portfolio rebalancing in case of a “multi-cycle” or “long-term” investor). Thus, in order to assess the potential for risk diversification across stock market cycles, it seems natural to consider correlations over the duration of a typical stock market cycle and not on, say, a daily or monthly basis. The latter correlation measures might offer a misleading view on the potential for risk diversification if investors base their rebalancing decisions on stock market cycles.

Also, persistent swings in stock market prices and the potentially destabilizing effects on the real economy raise the issue of how monetary authorities should respond. Indeed, bullish stock markets can induce large amounts of loan collateral - especially in less developed banking systems with poor regulatory frameworks - which then increase demand and goods price inflation. Moreover, when the stock market bulls turn into bears, this can result in widespread liquidity problems and a “credit crunch” in the financial system. Thus, monitoring the impact of stock market swings is also of potential interest to regulatory bodies caring about systemic risk and overall financial stability. Finally, if stock cycles have become more synchronized over time, the potential for financial system instability to spill over to other countries has also increased which suggests that a coordinated effort

---

of future fundamental values), see e.g. Summers (1986), Shiller (2000) or Anderson et al. (2003). These papers argue that prices can sometimes display seemingly persistent deviations from their long-run equilibrium values. Another view states that, although market sentiment can drive prices away from fundamentals in the short run, proportional differences between market prices and fundamentals are kept within bounds and stock markets exhibit a long-run relation between prices and fundamentals, see e.g. DeLong (1992), Siegel (1998) or Coakley and Fuertes (2006). Our research does not fit in either of these two strands of literature as we do not aim to disentangle the causes of stock market bulls and bears.

<sup>3</sup>Few empirical studies identified and investigated univariate features of stock market cycles, see e.g. Edwards et al. (2003), Gómez Biscarri and Pérez de Gracia (2003), Pagan and Sossounov (2003) or Lunde and Timmermann (2004). An even smaller set of papers looked into whether stock market cycles comove, see e.g. Gómez Biscarri and Pérez de Gracia (2003), Edwards et al. (2003) and Harding and Pagan (2006b).

of policymakers and regulatory bodies is necessary.

This chapter makes several contributions to the stock market bulls and bears cum synchronization literature.<sup>4</sup> More specifically, we extend the Generalized Method of Moments (GMM) approach to measuring business cycle synchronization due to Harding and Pagan (2006b) toward estimating and testing for (bivariate, multivariate) cyclical stock market synchronization. The present chapter builds further on their framework. First, we test the null hypothesis of perfect synchronization. If that null hypothesis is rejected, we test the null hypothesis of “imperfect synchronization”. This basically amounts to testing the cross equality of all bivariate binary correlations for all possible restricted synchronization values strictly smaller than 1 (including 0). This seems a more realistic null hypothesis than testing against either a value of 0 or 1 only. To this aim, we can use the same test statistic that Harding and Pagan proposed for testing the null hypothesis of “perfect nonsynchronization”: the limiting distribution stays the same as long as the value of the restricted synchronization index under the null hypothesis stays strictly below 1.

Moreover, and provided the null hypothesis of imperfect synchronization is not always rejected, the test for imperfect synchronization produces an estimator for the multivariate synchronization index. First, the test renders the range of GMM restricted synchronization estimates that do not lead to rejection of the cross-equality hypothesis for the bivariate correlations. Next, one can select the GMM estimate that minimizes the test statistic as “best attainable” (i.e., that leads to the strongest nonrejection) estimate for the “multivariate imperfect synchronization index”.

Also, we show by Monte Carlo simulation that the asymptotic version of the synchronization test is biased toward rejection and that the size distortion increases with the number of countries considered. A bootstrap procedure for the small sample critical values of the imperfect synchronization test is shown to remove the size distortion nearly entirely, provided the number of countries is not too big relative to the time series length of the cycle indicators.

The remainder of the chapter is organized as follows. In Section 3.2 the

---

<sup>4</sup>Preceding literature on stock market synchronization is vast and difficult to summarize. Early research on stock market linkages mainly documented cross border return correlations (see e.g. Roll (1988)). This correlation analysis was refined later on, either by implementing multivariate stochastic volatility (SV) models or cointegrated Vector Auto Regressions (VAR). Representative articles of the former “school” include King and Wadhwani (1990); Mallaris and Urrutia (1992); Lin et al. (1994); Susmel and Engle (1994). These ARCH-type models were used, inter alia, to investigate the direction of international spillovers as well as to identify differences in market comovements in periods of market turbulence and market quiescence. Baillie and Bollerslev (1989), however, argued that the modeling of returns can result in a loss of information on possible common trends when prices are cointegrated. Representative articles of the cointegrated VAR literature are Kasa (1992) and Click and Plummer (2005).

Harding-Pagan framework for measuring and testing cycle synchronization is shortly revisited. Section 3.3 contains a Monte Carlo investigation of the multivariate nonsynchronization (SMNS) test for varying numbers of countries and time series lengths. A bootstrap algorithm for size-corrected critical values is proposed in the same section. Proofs for the consistency of the bootstrap procedure are provided in Section 3.4. The empirical implementation we show in Section 3.5. Conclusions are drawn in Section 3.6.

## 3.2 Strong multivariate synchronization

Harding and Pagan (2006b) introduced the concept of synchronization for measuring business cycle comovements. We will brief their GMM framework here.

Prior to calculating cyclical correlations we need to identify what stock market “bulls” and “bears” are. The financial press nowadays usually focuses on increases (declines) of the market being greater (less) than either 20% or 25%, see Pagan and Sossounov (2003). As for the academic literature, there is no consensus on what bulls and bears actually mean. One definition describes bull or bear markets as “periods of generally increasing/decreasing market prices”, see Chauvet and Potter (2000, p.90, fn.6). The former definition, by focusing on extreme movements, would be analogous to “booms” and “busts” in the real economy whereas the latter definition seems closer to reflecting business cycle contractions and expansions. We use the latter definition of bulls and bears that focuses on how stock prices evolve between local peaks and troughs. This approach is in line with the business cycle literature going back to Burns and Mitchell (1946). The definition essentially implies that a bullish stock market turned bearish if prices have declined for a substantial period since their previous (local) peak. Such a definition does not exclude sequences of price falls (rises) during a bull (bear) phase, but there are restrictions on the extent to which these sequences of price reversals can occur and yet still be considered part of any given bull or bear phase.

Let  $p_{it}$  denote the log stock price for country  $i$  at time  $t$  ( $i = 1, \dots, n; t = 1, \dots, T$ ). Bull and bear periods are determined using the marginal transform  $\varphi(\cdot)$  such that  $\varphi(p_{it}) = S_{it}$  ( $\forall i$ ) where  $S_{it}$  is 0 or 1 in case of bear or bull period, respectively. There are two main methodological strands in the literature to select  $\varphi(\cdot)$ . First, Hamilton (1989) imposes a two regime Markov-switching model on  $\varphi(\cdot)$  that allows for booms and busts.<sup>5</sup> We prefer the second, nonparametric approach which can be motivated by the complex temporal behavior of financial time series (clusters of volatility, long memory etc.) and the resulting risk of model

---

<sup>5</sup>Applications of this approach include Hamilton and Lin (1996) and Maheu and McCurdy (2000).

misspecification. The key feature of nonparametric filters or dating algorithms is the location of turning points (peaks and troughs) that correspond to local maxima and minima of the series. Loosely speaking, a peak/trough in the (log) stock price series  $p_{it}$  occurs when  $p_{it}$  reaches a local maximum (minimum) in a window of six months width.<sup>6</sup> For a peak to occur at  $t$ , this implies:

$$p_{i,t-6} \cdots p_{i,t-1} < p_{i,t} > p_{i,t+1} \cdots p_{i,t+6}.$$

Likewise there will be a trough at  $t$  if

$$p_{i,t-6} \cdots p_{i,t-1} > p_{i,t} < p_{i,t+1} \cdots p_{i,t+6}.$$

The above dating algorithm constitutes the core of a more complicated method proposed by Bry and Boschan (1971) in order to date business cycle phases at the quarterly frequency. This approach has since also been used to determine stock market cycles, see e.g. Pagan and Sossounov (2003) or Kaminski and Schmukler (2008). Once turning points are determined, bull and bear periods can be identified as periods between troughs (peaks) and peaks (troughs), respectively.<sup>7</sup>

Harding and Pagan (2006b) proposed GMM-based multivariate procedures for testing the null hypothesis that business cycles are either perfectly synchronized or not synchronized at all. These two borderline null hypotheses might be a bit unrealistic. So, we propose to test the “weaker” null hypothesis that  $\rho^{12} = \dots = \rho^{n(n-1)} = \rho_0$  with  $-1 < \rho_0 < 1$ , where  $\rho^{ij}$  is the cross-correlation of country  $i$ ’s and country  $j$ ’s bull and bear markets. We call this null hypothesis Strong Multivariate Synchronization of order  $\rho_0$  ( $SMS(\rho_0)$ ) throughout the rest of the chapter.<sup>8</sup> So,  $\rho_0$  constitutes the “common” correlation or synchronization index shared by all countries. The larger  $\rho_0$  the stronger multivariate synchronization is supposed to be. In order to test this hypothesis one can use the same framework introduced in Harding and Pagan (2006b). Their procedure starts from the following  $n(n+1)/2$  moment conditions under the null hypothesis  $SMS(\rho_0)$ :

$$E[h_t(\theta_0, S_t)] = 0 \tag{3.1}$$

---

<sup>6</sup>Robustness checks for different window widths are not included for sake of space considerations. Turning point locations only change marginally.

<sup>7</sup>Throughout the rest of the paper we allow for two regimes, but generalizing the framework is straightforward.

<sup>8</sup>The borderline case of perfect multivariate synchronization ( $\rho_0 = \pm 1$ ) is excluded from the analysis because it is unlikely to be observed and the asymptotic distribution in this limiting case turns out to be a weighted average of  $\chi^2$  distributions with the weights to be determined by simulation, see Gouriéroux et al. (1982).



with

$$h_t(\theta_0, S_t) = \begin{bmatrix} S_{1t} - \mu_1 \\ \vdots \\ S_{nt} - \mu_n \\ \frac{(S_{1t} - \mu_1)'(S_{2t} - \mu_2)}{\sqrt{\mu_1(1-\mu_1)\mu_2(1-\mu_2)}} - \rho_0 \\ \vdots \\ \frac{(S_{(n-1)t} - \mu_{n-1})'(S_{nt} - \mu_n)}{\sqrt{\mu_{n-1}(1-\mu_{n-1})\mu_n(1-\mu_n)}} - \rho_0 \end{bmatrix} \quad (3.2)$$

and with  $\theta'_0 = [\mu_1, \dots, \mu_n, \rho_0, \dots, \rho_0]$ , the restricted vector under the  $SMS(\rho_0)$  case. The expectations operator in (3.1) is defined over the time dimension. The first subset of  $n$  moment conditions in (3.1) and (4.2) defines the population means of the cycle dummies and reflects the likelihoods of the stock markets to be in the bullish phase. The remaining  $n(n-1)/2$  moment conditions express equality of all cyclical correlations to some common value  $\rho_0$ . If the above moment conditions hold, stock markets, albeit imperfectly synchronized, do exhibit a “common” or “homogeneous” synchronization index  $\rho_0$ .

We propose to test  $SMS(\rho_0)$  via Hansen’s (1982) Wald test statistic:

$$W(\rho_0) = Tg(\theta_0, [S]_{t=1}^T)' \hat{V}^{-1} g(\theta_0, [S]_{t=1}^T), \quad (3.3)$$

which converges to an asymptotic  $\chi^2_{n(n-1)/2}$  distribution under the null hypothesis  $\theta = \theta_0$ , see e.g. Harding and Pagan (2006b, p.70). The covariance matrix  $\hat{V}$  is a heteroscedasticity and autocorrelation consistent (HAC) estimator of the variance-covariance matrix of  $\sqrt{T}g(\theta_0, [S]_{t=1}^T)$ , see e.g. Newey and West (1987). The statistic is defined as a quadratic form in the penalty vector  $g(\theta_0, [S]_{t=1}^T) = \frac{1}{T} \sum_{t=1}^T h_t(\theta_0, S_t)$  which reflects the average deviation from the moment conditions. The stronger the deviations from the moment conditions the more likely a rejection of  $SMS(\rho_0)$  becomes.

In order to solve the problem of testing against an unknown value of  $\rho_0$ , we calculate the GMM test statistic in (3.3) for different values of  $\rho_0$  ( $-1 \leq \rho_0 < 1$ ). This grid search renders the interval of GMM estimates  $[\rho_-, \rho_+] \subset ]-1, 1[$  that do not lead to rejection of the null hypothesis  $SMS(\rho_0)$  at a pre-specified nominal size. From this interval estimate, we select the GMM point estimate  $\hat{\rho}_0$  that minimizes the test statistic  $W$ .<sup>9</sup> Or in formula form:

$$\hat{\rho}_0 = \underset{\rho \in [\rho_-, \rho_+]}{\text{Argmin}} \left\{ Tg(\theta_0, [S]_{t=1}^T)' \hat{V}^{-1} g(\theta_0, [S]_{t=1}^T) \right\}. \quad (3.4)$$

---

<sup>9</sup>More sophisticated optimization algorithms like e.g. the Newthton-Raphson technique could also be applied in order to determine the minimum point of the test statistic.

Before putting the estimator of the common synchronization index cum testing procedure at work in an empirical application we need to evaluate the small sample behavior of the test. It can be suspected that the asymptotic distribution of the W-test only poorly approximates the test's small sample behavior when the number of countries grows large.

### 3.3 Small sample properties of the $\text{SMS}(\rho_0)$ test

Previous papers (see e.g. Christiano and den Haan (1996) and Koenker and Machado (1999)) already argued that existing asymptotic theory for GMM estimators may break down in small samples. In this section we investigate whether the GMM-based asymptotic test for multivariate nonsynchronization (SMNS) in (3.3) also suffers from small sample problems. We suspect that this problem might be more severe when the number of countries (cross sectional dimension) grows large relative to the length of the time series.

We set up a Monte Carlo experiment for six different Data Generating Processes (DGPs) that we consider as sufficiently representative for the current stock market cycle literature. First, we draw  $(n \times 1)$  vectors  $X_t$  ( $t = 1, \dots, T$ ) from a multivariate standard normal distribution with unit marginal variance and equal pairwise correlation  $\rho$ . The normal draws under  $DGP_1$  can be interpreted as stationary fluctuations around a trend (sometimes also called a “growth” cycle, see e.g. Burns and Mitchell (1946) for an early reference). The second DGP models the cycle as a simple random walk without a drift,  $\Delta X_t = u_t$ . The disturbance vector  $u_t$  is drawn from a multivariate normal df conform with  $DGP_1$ . Notice that  $DGP_2$  possesses a stochastic trend without drift. If one considers the growth cycles, i.e. removing the stochastic trend, they are strictly identical to those obtained in  $DGP_1$ . Clearly the distinguishing feature between the two DGPs lies in the stationarity properties of the generated series.

$DGP_3$  also considers the cycles in levels but contrary to  $DGP_2$  it allows for a constant drift.  $DGP_3$  is more realistic and corresponds to the following  $ARI(1)$  process:

$$\Delta X_t = 0.0042 + 0.435\Delta X_{t-1} + u_t,$$

with  $u_t \rightarrow N(0, 0.016)$ . The series in levels ( $X_t$ ) are built following the rule  $X_t = X_{t-1} + \Delta X_t$  assuming an initial value  $X_0 = 0$ . Similarly,  $DGP_4$  exhibits a constant drift but instead of being autoregressive it is characterized by a moving average ( $IMA(1)$ ). It has the following form:

$$\Delta X_t = 0.0074 + u_t + 0.596u_{t-1},$$

with  $u_t \rightarrow N(0, 0.015)$ . The series in levels are built following the similar rule as for  $DGP_3$ .<sup>10</sup>

The last two remaining DGPs are as follows.  $DGP_5$  corresponds to the first difference of  $DGP_3$  and  $DGP_6$  to the first difference of  $DGP_4$ . Both are thereby stationary.

Next to generating raw data in different ways, we also consider two algorithms for dating the cycle phases. First, the “calculus rule” attaches a one to positive random draws and a zero otherwise. As it requires to work on growth cycles, it can only be applied to  $DGP_s$ , which are stationary around a trend (stochastic or deterministic). The second filter is the Bry and Boschan (1971) dating algorithm, which we already outlined above. Once turning points are determined, bull and bear periods can be identified as periods between troughs (peaks) and peaks (troughs), respectively.<sup>11</sup>

The Monte Carlo investigation combines different DGPs with different dating algorithms. First, the series in levels generated by the  $DGP_s$  are associated with BBQ. As  $DGP_1$  is stationary by definition, it can be coupled with the calculus rule, which leads to the same results as the series generated by  $DGP_2$  once its stochastic trend is removed by taking the first difference ( $\Delta Y_t$ ).  $DGP_3$  and  $DGP_4$  can only be filtered with the BBQ method as they are trending. Finally, the series generated by  $DGP_5$  and  $DGP_6$ , are coupled with the calculus rule and BBQ.

Without loss of generality we limit ourselves to analyzing the size distortion of (3.3) under the null hypothesis of strong multivariate non-synchronization ( $SMNS$ ) (i.e.  $\rho^{ij} = 0$ , for  $i \neq j$ ).<sup>12</sup>

Figure 3.1 shows the small sample size (nominal size equal to 5%) of the asymptotic test for  $SMNS$  ( $\rho = 0$ ) as a function of  $n$  (the number of countries) and  $T$  (the length of the time series). The horizontal axes allow for an upper bound of ten countries and 1,000 time series observations. First, the outcomes do not seem to differ greatly across different DGPs or dating algorithms. More importantly, however, the rejection rates reveal that size distortion grows rapidly with the number of countries and is only negligible in the bivariate case (first row). In the worst case scenario of ten countries the asymptotic GMM test nearly always rejects the null hypothesis of absence of synchronization, even with time series of

---

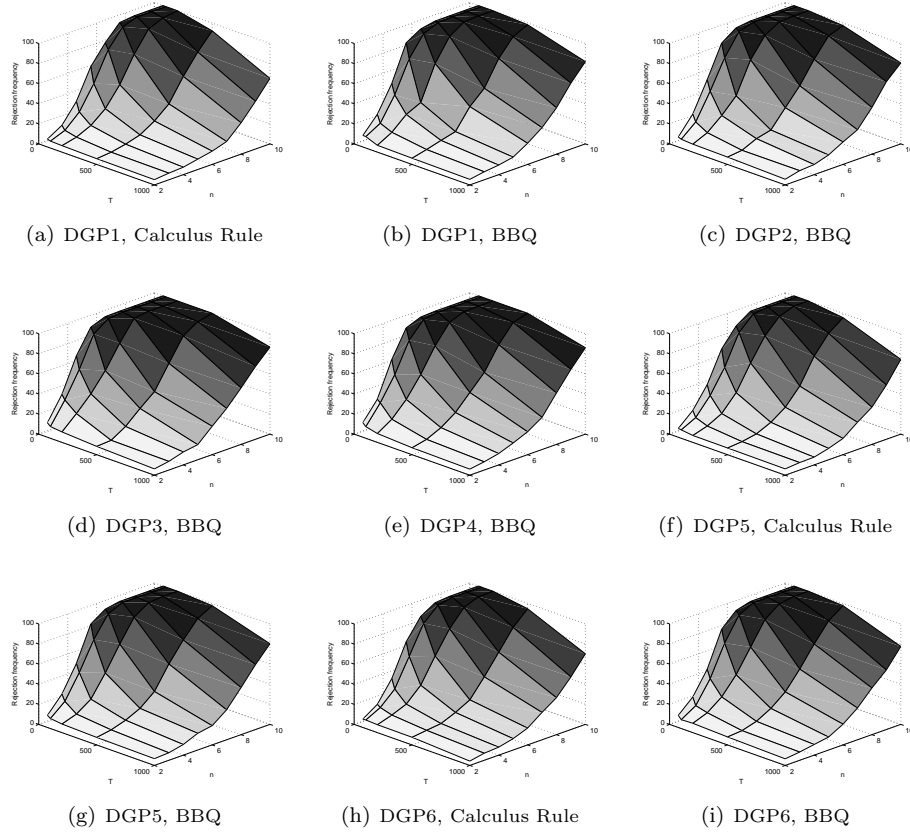
<sup>10</sup>The parameters used here are obtained from fitting the respective models on industrial production index data from the US from 1970 until 2007. Such an application is based on the paper Candelon et al. (2008a), and shall only serve as an illustration of the finite sample size and power properties of the test. In Chapter 4 we also use more realistic DGPs for financial data and arrive essentially at the same conclusions.

<sup>11</sup>Throughout the rest of the chapter we allow for two regimes, but generalizing the framework is straightforward.

<sup>12</sup>Similar conclusions on small sample behavior hold when simulating under the null hypothesis  $SMS(\rho_0)$ , with  $\rho_0 \neq 0$ . The latter simulations are therefore omitted for sake of space considerations.

1,000 observations. To better grasp the intuition behind this outcome, notice that the number of moment conditions to be estimated in (3.1) using time series of fixed length  $T$  grows more rapidly than the number of countries  $n$ . For example,  $\frac{10 \times 9}{2} = 45$  moment conditions have to be estimated for a panel of  $n = 10$  countries.

Figure 3.1: Size of the asymptotic  $\text{SMS}(\rho_0)$  test

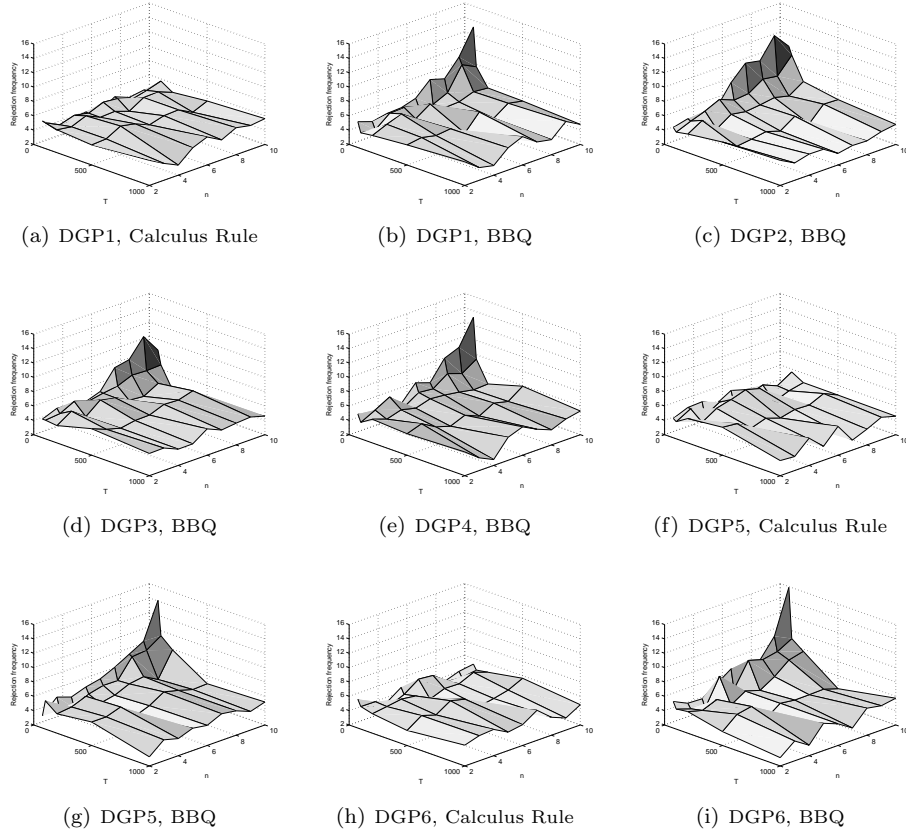


**Note:** Z-axis indicates the rejection frequency of the asymptotic version of  $\text{SMS}(\rho_0)$  test under several DGP and dating rules for business cycles (indicated below the figure). T- and n-axis report respectively the sample size  $T$  and the number of countries  $n$ .

Lengthening the time series removes at least part of the problem. Indeed, Figure 3.1 shows that the size distortion somewhat decreases for larger values of  $T$  but far from sufficient to talk about a “size corrected” test. Additional Monte Carlo simulations for the 10 country “worst case” revealed that one would need time series of at last 10,000 data points in order to have a nondistorted test statistic. However, these sample lengths are nonrealistic in financial cycle

research when one uses monthly data. We therefore propose to bootstrap the small sample distribution of the GMM-test. The asymptotic pivotality of the W-test (3.3) and its continuity in the parameter vector  $\theta_0$  ensure proper convergence of the bootstrapped df to the true asymptotic  $\chi^2_{(n-1)n/2}$  distribution, see Section 3.4 of this chapter and, for example, Horowitz (2001).

Figure 3.2: Size of the bootstrap SMS( $\rho_0$ ) test



**Note:** Z-axis indicates the rejection frequency of the bootstrap version of SMS( $\rho_0$ ) test under several DGP and dating rules for business cycles (indicated below the figure). T- and n-axis report respectively the sample size  $T$  and the number of countries  $n$ .

Apart from asymptotic convergence, we also want to know whether the bootstrapped critical values are sufficiently size-corrected in small samples. Analogous to the preceding size study for the asymptotic test, we study (without loss of generality) the performance of the bootstrap under the null hypothesis of Strong Multivariate Nonsynchronization (SMNS), i.e.,  $H_0 : \rho_0 = 0$ .

The bootstrap algorithm that we implement is (semi-)parametric in nature (only the block length parameter has to be chosen). Obviously, opting for either a parametric or (semi-)parametric bootstrap is to some extent a matter of taste and it is impossible to claim that one approach is strictly better than the other. Our viewpoint is that it is rather unrealistic to assume that researchers know the exact model of the cycle. Moreover, not imposing parametric cycle models prevents the risk of misspecification. Also, popular dating algorithms like the Bry-Boschan (BBQ) algorithm are semi-parametric (the window width being the only parameter to be chosen).

The bootstrap is performed in blocks to account for the temporal persistence in 0-1 dummies describing the cycle phases. There is no tailor-made solution to the determination of the optimal block length  $bl_{opt}$ . Hall et al. (1995) suggest to choose  $bl_{opt} = c.T^{1/4}$ . However, the parameter  $c$  is only known for certain specific parametric specifications and we just argued that we do not want to impose parametric cycle models and dating procedures. We therefore opt for a heuristic approach (see step BO2 below): we iteratively determine the block length such that the average cycle duration in the original sample is preserved in the bootstrapped samples.

In order to evaluate the bootstrap's small sample performance we implement the following multi-stage Monte Carlo cum bootstrap procedure:

1. Step MC1: Consider a particular DGP and a dating algorithm  $F(\cdot)$ . Draw a  $T \times n$  matrix of standard normally distributed residuals  $U$  with characteristic element  $u_{i,t}^{MC}$ . Use the residuals in order to build  $n$  time series  $X_{i,t}^{MC}$ , with  $t = 1, \dots, T$  and  $i = 1, \dots, n$ . This renders the  $T \times n$  matrix  $X$ .
2. Step MC2: Use the dating algorithms to build a  $T \times n$  binary variable matrix  $S$  with characteristic element  $F(X_{i,t}^{MC}) = S_{i,t}^{MC}$ . Calculate the cross sectional average number of cycle phases  $bc$ .
3. Step MC3: Compute the simulated W-statistic for  $\rho_0 = 0$ , i.e.,  $W^{MC}(\rho_0 = 0)$  using  $S$  as input.
4. Step BO1: Generate a  $T \times n$  matrix of bootstrap replications  $S^B$  by randomly drawing 25 consecutive  $S_t$  with replacement.
5. Step BO2: Calculate the average number of cycle phases in the bootstrapped sample  $bc_{boo}$ . If  $bc_{boo} > 1.05bc$  (resp.  $bc_{boo} < 0.95bc$ ), increase (resp. decrease) by 5 the consecutive  $S_t$  which are drawn in BO1. Repeat BO1-BO2, until  $bc_{boo} \in [0.95bc, 1.05bc]$ .<sup>13</sup>

---

<sup>13</sup>Robustness checks with  $[0.975bc, 1.025bc]$  have been performed and provide similar results.

6. Step BO3: Compute the bootstrapped W-statistic for  $\rho_0 = 0$ , i.e.,  $W^B(\rho_0 = 0)$  using  $S^B$  as input.
7. Repeat the bootstrap steps 4-6 a “sufficient” number of times ( $M$ ).  $M$  is endogenously determined using the three step-method of Andrews and Buchinsky (2000).<sup>14</sup> A critical value is then obtained as the  $\alpha$  percent quantile, say  $W_{crit}^B$ , from the empirical distribution of the bootstrap test statistic. The nominal size is set to  $\alpha = 5\%$ . The null hypothesis of  $SMNS(\rho_0 = 0)$  is rejected if  $W^{MC}(\rho_0 = 0) > W_{crit}^B$ .
8. Repeat steps 1-7 a large number of times to obtain the rejection frequencies of the test  $W(\rho_0 = 0)$ .

Figure 3.2 summarizes the results of the Monte-Carlo experiments and shows the rejection rates of the bootstrap-based version of the test for SMNS ( $\rho = 0$ ) as a function of  $n$  (the number of countries) and  $T$  (the length of the time series). We consider the same combinations of *DGP* and dating algorithm as in the previous figure and the nominal size is again set equal to 5%.

It turns out that the size distortion is greatly reduced as compared to the asymptotic version of the test for most combinations of  $T$  and  $n$ . However, even the bootstrap no longer seems to be a valid remedy for the size distortion in case  $T$  is small relative to  $n$ . For example, in the case of a sample size  $T = 75$ , considering a number of countries  $n$  larger than six would still lead to overrejection of the null hypothesis of SMNS ( $\rho = 0$ ). Thus, this highlights the limits of our procedure when information is restricted, i.e.  $T$  is relatively small with respect to  $n$ .

Figure 3.2 also reveals that size distortions are negligible when cycles are dated using the calculus rule. Our intuition is that distortion also depends on the relative cycle length ( $bc$ ) compared to the sample size ( $T$ ). We expect that when the average length of the cycles will be low (i.e. more numerous cycle phases), the test will converge quicker to the asymptotic distribution. To investigate this point, the link between the size distortion and the average length of the cycle across our nine experiments is explored using the asymptotic and the bootstrap versions of the test. If our intuition is correct then a positive relationship between the average cycle length and the size distortion should appear. In Figure 3.3, we report all the size distortions and the average lengths of cycles associated with our nine experiments in the case of  $(n = 5, T = 200)$  and  $(n = 10, T = 75)$ .<sup>15</sup>

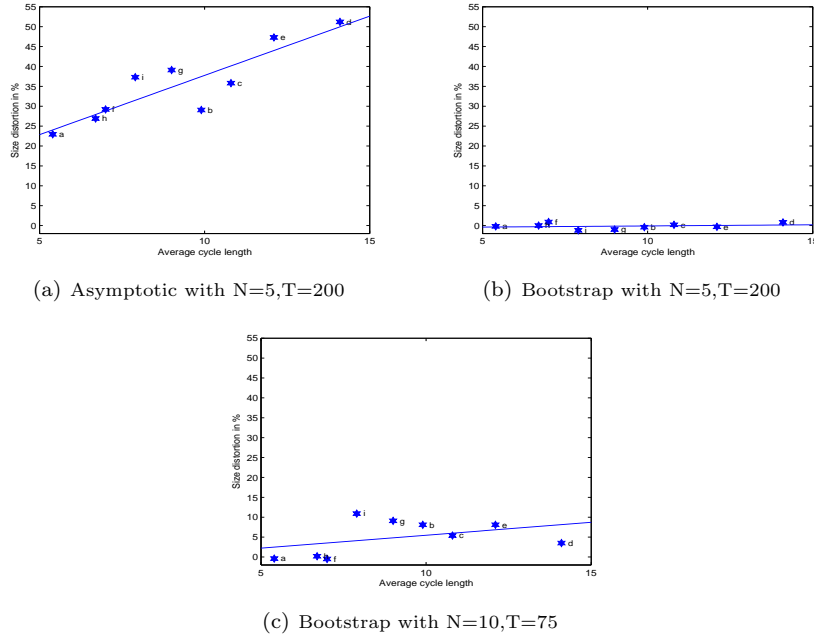
---

<sup>14</sup>The endogenous number of bootstrap replications  $M$  was found to vary between 200 and 2,000.

<sup>15</sup>In order to save space, we only report these two cases, but the positive link between the size distortion and the cycle length appears to be consistent in all other cases. Tables can be obtained from the authors upon request.

First it appears that the calculus rule leads in general to a smaller cycle length than the BBQ (using the calculus rule business cycles exhibit an average length of 6.36 periods versus 10.63 periods with BBQ). Second, it turns out that the positive relationship between the cycle length and size distortion is supported by our experiments for the asymptotic version of the test. Third, the relationship is no more present when using the bootstrap version of the test for  $(n = 5, T = 200)$ , but still supported for  $(n = 10, T = 75)$ . From a theoretical point of view, this result indicates that the convergence speed of the bootstrap version of the test<sup>16</sup> depends on  $T$  and  $n$  but also on  $bc$ . It is straightforward to notice that if  $T \rightarrow \infty$  then  $bc/T \rightarrow 0$ . So, asymptotically  $bc$  does not affect the convergence properties. But when  $T$  is small then the speed of convergence can be higher if the cycle length  $bc$  is short.

Figure 3.3: Size distortion and average cycle length



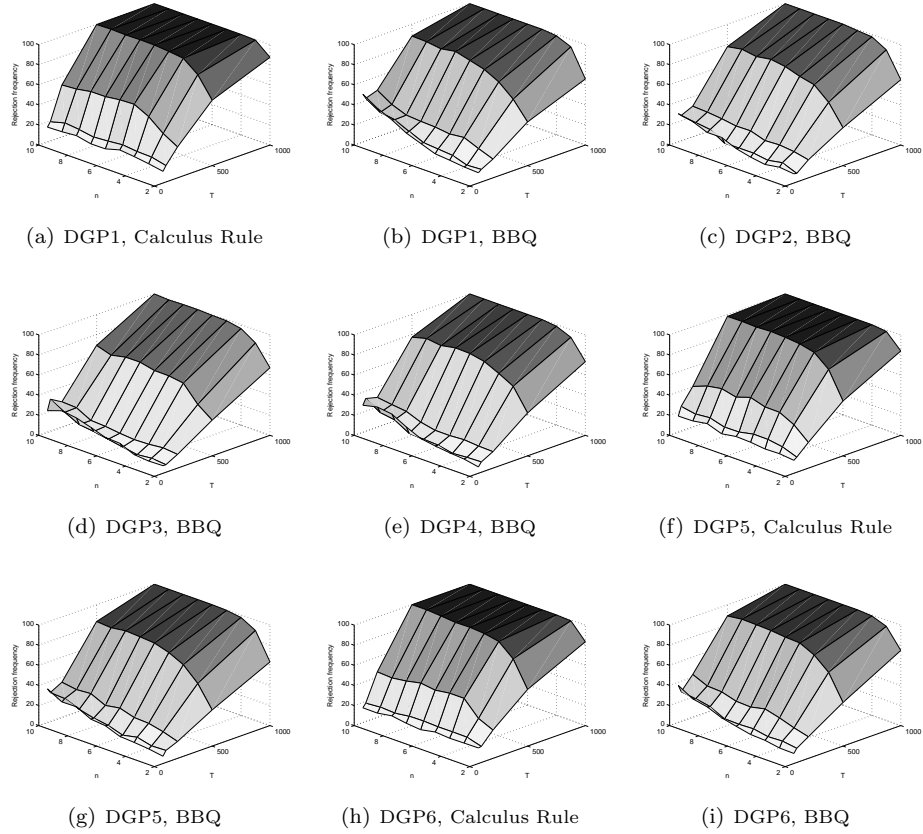
**Note:** Stars indicate the different DGP and filter combinations with their respective average cycle lengths and size distortions. The lines represent the fitted OLS regression lines.

From an empirical perspective it is clear that the choice of the filter either being

<sup>16</sup>A formal derivation of this convergence properties is beyond the scope of this chapter but leaves us with a very interesting topic for future research. Nevertheless, Figure 3.3 provides the reader with a good intuition of the convergence speed.



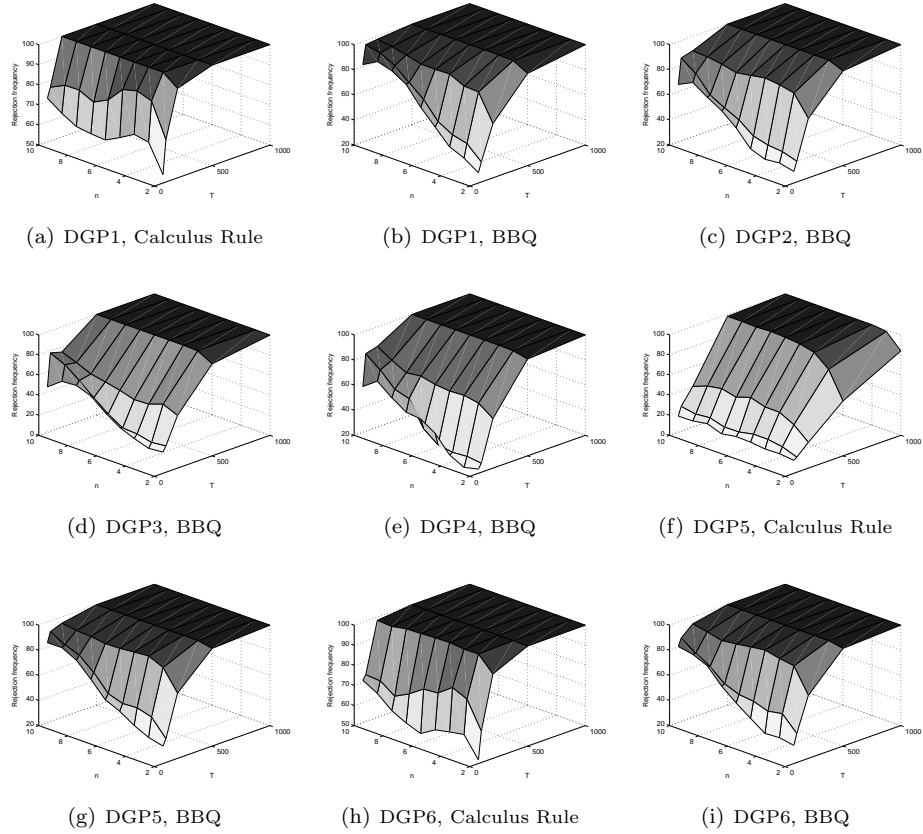
Figure 3.4: Power of the bootstrap SMS( $\rho_0$ ) test against  $\rho = 0.10$



**Note:** Z-axis indicates the rejection frequency of the bootstrap version of SMS( $\rho_0$ ) test under several DGP, exhibiting SMS( $\rho_{0.1}$ ) and dating rules for business cycles (indicated below the figure). T- and n-axis report respectively the sample size  $T$  and the number of countries  $n$ .

### 3.3 SMALL SAMPLE PROPERTIES OF THE $\text{SMS}(\rho_0)$ TEST

Figure 3.5: Power of the bootstrap  $\text{SMS}(\rho_0)$  test against  $\rho = 0.25$



**Note:** Z-axis indicates the rejection frequency of the bootstrap version of  $\text{SMS}(\rho_0)$  test under several DGP, exhibiting  $\text{SMS}(\rho_{0.25})$  and dating rules for business cycles (indicated below the figure). T- and n-axis report respectively the sample size  $T$  and the number of countries  $n$ .

the “calculus rule” or “BBQ” to a great part depends on what the study wants to focus on. For example, if the researcher would like to analyze the “growth cycle” based on detrended series he should use the calculus rule because BBQ would not make sense here. But if the classical cycle based on some series in levels shall be considered the calculus rule has to be replaced by the BBQ filter because of the non-stationarity of the underlying series. In our simulation study we can work with both filters on different series because the trend properties are perfectly known. In the empirical application, though, we will focus on different series in levels and on the classical cycle. Therefore, we only use the BBQ dating algorithm. This is in line with many studies in particular the companion paper of Harding and Pagan (2006b).<sup>17</sup>

Finally, we also investigated the small sample power of the bootstrap-based W-test. The power is size-corrected in the sense that we used the bootstrapped critical values obtained from our algorithm. Power results are summarized in Figure 3.4 and 3.5.

The power is calculated under the null hypothesis of strong Multivariate Non-synchronization (SMNS) against the alternatives  $\rho = 0.10$  and  $\rho = 0.25$ . We find that the power is of acceptable magnitude, even for low values of  $T$  and  $n$ . Moreover, the power does not differ much across different DGPs and dating algorithms. As expected, the power we find to rise with the sample size  $T$ .<sup>18</sup>

### 3.4 Consistency of the block bootstrap method

In this section we show that the proposed block bootstrap method indeed is consistent and can be used in this setting. In his seminal paper Hansen (1982) derived the large sample properties of generalized method of moment (GMM) estimators. He shows that GMM estimators are consistent and asymptotically normally distributed if the DGP and corresponding moment conditions fulfill certain regularity conditions. In order to show the consistency of our block-bootstrap version of the Harding and Pagan (2006b) GMM test we take the following route. First, we present the by now well-known conditions and assumptions important for the asymptotic consistency and normality of the GMM estimator as they were derived in Hansen (1982). Realizing that those results only hold asymptotically and may not be good approximations in finite samples (as in our case), we propose a bootstrap version of the Harding and Pagan (2006b) test which better approximates

---

<sup>17</sup>A comparison of classical and growth cycle results using our proposed methodology would be very interesting but is left for further research.

<sup>18</sup>The power also rises when  $T$  is small relative to  $n$  which seems rather counterintuitive. However, given the earlier observation that the bootstrap is unable to mitigate the size distortion in this situation, the size-corrected power results cannot be trusted.

the finite sample distribution of such a test. Such a bootstrap procedure also asks for some conditions to be satisfied by the DGP and functionals leading to the test statistic. Those conditions we can find in the works of, for example, Hall and Horowitz (1996), Lahiri (1999), and Haerdle et al. (2003). As a next step we show that the DGPs of our series in the simulation and the empirical section are very likely to fulfill such conditions setting the ground for the validity of our block-bootstrap procedure.

We start with presenting the conditions for the GMM estimator to converge to the parameter vector that is being estimated and for its asymptotic normal distribution. Most of the following will be taken from Hansen (1982). Define the  $n$  component stochastic process  $\{x_t : t \geq 1\}$ , that is defined on the probability space. The observable data series can be thought of as a finite segment of the realization of such a stochastic process  $\{x_t(\omega_0) : 1 \leq t \leq T\}$ , where  $T$  stands for the sample size and some  $\omega_0 \in \Omega$  represents one realization of the stochastic process drawn from  $\Omega$ . Let  $C$  be a parameter space that is a subset of  $R^q$  and let  $\theta_0$  be the element of  $C$  we wish to estimate. From an econometrical model there emerge  $r$  orthogonality conditions that can be expressed in the function  $h$ . For simplicity we assume that  $r = q$ .<sup>19</sup> For our GMM estimator we need the following assumptions for consistency and asymptotic normality:

ASSUMPTION 3.1: The stochastic process  $\{x_t : t \geq 1\}$  is stationary and ergodic.

ASSUMPTION 3.2:  $(C, \sigma)$  is a separable metric space, where  $\sigma$  can be defined with some absolute value norm on  $R^q$  that contains  $\theta_0$ .

ASSUMPTION 3.3:  $h(\cdot, \theta)$  and  $\partial h / \partial \theta(x, \cdot)$  are Borel measurable for each  $\theta$  in  $C$  and  $h(x, \cdot)$  and  $\partial h / \partial \theta(x, \cdot)$  are continuous on  $C$  for each  $x$  in  $R^q$ .

ASSUMPTION 3.4:  $Eh(x_1, \theta)$  exists and is finite for all  $\theta \in C$  and  $Eh(x_1, \theta_0) = 0$ .

With a method of moments estimator of  $E[h(x_t, \theta)]$  we obtain an objective function whose minimizer is the estimator of  $\theta_0$ . Let

$$\begin{aligned} h_t(\omega, \theta) &= f[x_t(\omega), \theta], \\ g_t(\omega, \theta) &= \frac{1}{T} \sum_1^T f[x_t(\omega), \theta], \\ B_T(\omega) &= \left\{ \theta \in C : |h_t(\omega, \theta)|^2 = \inf_{\theta \in C} |h_t(\omega, \theta)|^2 \right\}. \end{aligned}$$

Then we can define:

DEFINITION 3.1: The GMM estimator  $b_t : T \geq 1$  is a sequence of random

---

<sup>19</sup>We can also allow for over-identifying conditions if we have  $r \geq q$ . This would ask for a weighing matrix, which Hansen (1982) calls  $a(\omega)$ . Such an extension is possible but does not add much at this stage.

vectors such that  $b_t(\omega) \in B_T(\omega)$  for  $T \geq T^*(\omega)$  where  $T^*(\omega)$  is less than infinity for almost all  $\omega$  in  $\Omega$ .

Under these assumptions we can write the following (see Hansen (1982) Theorem 2.1):

**THEOREM 3.1:** Suppose Assumptions 3.1-3.4 are satisfied. If (i)  $h_1$  is first moment continuous for all  $\theta \in C$ ; (ii)  $C$  is compact; (iii)  $h_0(\theta)$  has a unique zero at  $\theta_0$ ; then a GMM estimator  $\{b_T : T \geq 1\}$  exists and converges almost surely to  $\theta_0$ .<sup>20</sup>

For asymptotic normality we need some further assumptions, which are also taken from Hansen (1982).

**ASSUMPTION 3.5:**  $\partial h_1 / \partial \theta$  is first moment continuous at  $\theta_0$  and  $E[d_0]$  where  $d_0 = \partial h / \partial \theta(x_1, \theta_0)$  exists, is finite, and has full rank.

Now define:

$$w_t = h(x_t, \theta_0)$$

and

$$v_j = E[w_0 | w_{-j}, w_{-j-1}, \dots] - E[w_0 | w_{-j-1}, w_{-j-2}, \dots] \quad \text{for } j \geq 0.$$

We can use an iterated expectations argument to show that  $\{v_j : j \geq 0\}$  is a martingale difference sequence. Using this we only need one more assumption:

**ASSUMPTION 3.6:**  $E[w_0 w'_0]$  exists and is finite,  $E[w_0 | w_{-j}, w_{-j-1}, \dots]$  converges in mean square to zero, and  $\sum_{j=0}^{\infty} E[v_j v'_j]^{1/2}$  is finite.

One can use Assumption 3.6 as a sufficient condition for applying a central limit theorem for stationary and ergodic processes as assumed in Assumption 3.1.<sup>21</sup> As a last step one can define the covariance matrices as:

$$\Gamma_w(j) = E[w_0 w'_{-j}].$$

With Assumptions 3.1 and 3.6 we know that  $\Gamma_w(j)$  is finite and that the following matrix is well defined:

$$V_w = \sum_{j=-\infty}^{+\infty} \Gamma_w(j).$$

With these assumptions and results at hand one can state the following:<sup>22</sup>

**THEOREM 3.2:** Suppose Assumptions 3.1-3.6 are satisfied. Then  $\{\sqrt{T}(b_T - \theta_0) :$

---

<sup>20</sup>For proofs see, for example, Hall (2005).

<sup>21</sup>For proofs please see, for example, Gordin (1969) and Hannan (1970).

<sup>22</sup>See Theorem 3.1 in Hansen (1982).

$T \geq 1\}$  converges in distribution to a normally distributed random vector with mean zero and covariance matrix  $d_0^{-1}V_w d_0^{-1'}$ .

This completes the asymptotic consistency and normality results for the GMM estimator. In order to be able to conduct block bootstrap simulations for the finite sample properties of the GMM estimators we have to make one more assumption about the stochastic process  $\{x_t(\omega)\}$  which can be found in Hall and Horowitz (1996). Other very good sources about the block bootstrap are, for example, Lahiri (1999) and Haerdle et al. (2003) where similar conditions can be found. Hall and Horowitz (1996, p.900) in their Assumption 1 assume a strong mixing condition for the block bootstrap procedure to yield asymptotic refinements:

ASSUMPTION 3.7: There is a sequence of i.i.d. vectors  $\{\epsilon_i : i = -\infty, \dots, \infty\}$  of dimension  $L_\epsilon \geq L_x$  and a  $L_x \times 1$  function  $f$  such that  $x_i = f(\epsilon_i, \epsilon_{i-1}, \epsilon_{i-2}, \dots)$ . There is a constant  $c > 0$  such that for all  $t = 1, 2, \dots$  and all  $m > d^{-1}$

$$\|f(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots) - f(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-m}, 0, 0, \dots)\| \leq d^{-1} \exp(-dm),$$

where  $\|\cdot\|$  is the Euclidean norm, the lag length  $m \rightarrow \infty$ , and  $d$  is some small number. Assumption 3.7 is a strong mixing assumption which basically implies that the underlying stochastic process “forgets” its past sufficiently fast. The mixing parameter is an exponentially decreasing function of the lag length.

In order to transfer this to binary variables as in our case we only need to establish that the corresponding DGP fulfills the assumptions above, namely that it is stationary, ergodic (Assumption 3.1), and strongly mixing (Assumption 3.7). If we take the example that  $\Delta y_t$  be a mean-zero stationary Gaussian process and combine this with the so-called “calculus rule” filter<sup>23</sup> we get a deterministic relationship between  $\Delta y_t$  and the binary transform  $S_t$ . Hannan (1980) and Harding and Pagan (2006a) show that:

$$\rho_S(k) = \frac{2}{\pi} \arcsin(\rho_{\Delta y}(k)),$$

where  $\rho_{\Delta y}(k) = \text{corr}(\Delta y_t, \Delta y_{t-k})$ , and  $\rho_S(k) = \text{corr}(S_t, S_{t-k})$ . Harding and Pagan (2006a) state that the DGP of constructed binary variables be determined by the interaction of the underlying stochastic process and the type of rule used for mapping the observable variables into the binary variables. They further find that such a DGP can generally be classified as a high-order Markov process that we often can approximate well with a first order process. Such a process can be shown to obey the rules and assumptions mentioned above. For example, Ibragimov and Linnik (1971), Rosenblatt (1971), and Jones (2004) show that ergodic

<sup>23</sup>Such a rule works as follows:  $\{S_t = 1 : \Delta y_t \geq 0\}$  and  $\{S_t = 0 : \Delta y_t < 0\} \forall t$ .

Markov chains are strongly mixing.<sup>24</sup> Therefore, the necessary assumptions are very likely to be satisfied and we can use the block bootstrap procedure in order to approximate the finite sample properties of the GMM estimator.

### 3.5 Empirical application

In this section we apply the multivariate synchronization measure to Asian stock markets. We estimate full sample bivariate and multivariate synchronization indices for the stock market cycles of five East Asian countries. We also distinguish full sample and subsample results to identify temporal shifts in synchronization. The choice for Asian countries can be motivated by the fact that changes in the synchronization of stock market cycles are probably more likely to occur across emerging markets than across developed markets due to financial liberalization and recurrent financial turmoil. In Chapter 4 we extend this framework with a stability test in order to test for changes in multivariate synchronization through time. There we detect a break around 1997. In this chapter we take this as an exogenously determined point to check the behavior of the series before and after 1997. In Chapter 4 we make the break point detection endogenous. Here we run subsample homogeneity tests in order to determine whether the multivariate synchronization (homogeneity) hypothesis breaks down in subsamples. Non-rejection of the homogeneity hypothesis could be interpreted as evidence for “common” stock market cycles. If the common cycle hypothesis does not hold for all considered stock markets, it might still be the case that it holds for a narrower subset, i.e., an Asian “core”. We therefore also pay attention to that possibility.

US dollar-denominated and dividend-adjusted monthly stock market indices for Singapore (S), Thailand (Th), South Korea (K), Taiwan (T) and Malaysia (M) were downloaded from the IFS database over the period January 1985 until November 2005 which amounts to 239 monthly observations. We selected those Asian countries with the longest possible time series in the database. As we are interested in measuring the comovement of medium-run fluctuations or cycles across stock markets, we first have to identify these “bulls” and “bears”. The cycle dummies are obtained by implementing the Bry and Boschan (1971) dating algorithm over a six month time interval. Previous studies suggest that stock market cycle dating and the resulting cyclical comovement measures are relatively robust across different types of (parametric, nonparametric) dating algorithms, see e.g. Candelon et al. (2008b). Figure 3.6 contains the evolution of the (log) stock indices for each of the countries, where the bull periods have been shaded to

---

<sup>24</sup>Also De Jong and Woutersen (2007) show that under some conditions a sequence of 0/1-valued random variables is strongly mixing with some mixing coefficient  $\alpha$ .

facilitate visual inspection.<sup>25</sup>

It is striking to see that the first half of the sample almost primarily consists of bull periods which illustrates why investors have been talking about “Asian Tigers” for a long time. Also, the 1997 Asian crisis and its aftermath are clearly visible in the plots. More specifically, notice that our results replicate the earlier finding that financial crises seem to erupt several months into bear phases (and sometimes very close to the end), see e.g. Edwards et al. (2003).<sup>26</sup> However, and as Edwards et al. (2003, p.936) pointed out, it would be premature to conclude on the basis of this observation that bear markets are leading indicators of financial crises. Somewhat surprisingly, the dotcom bubble burst is also clearly visible despite the relative underrepresentation of technology companies in emerging markets. Last but not least, the figures provide casual evidence for comovement or “synchronization” between bull and bear periods across markets. In order to assess the degree of synchronization we have to resort to the more advanced statistical tools introduced in the previous sections.

The cycle dummies  $S_{i,t}$  ( $i = 1, \dots, n; t = 1, \dots, T$ ) that result from applying the Bry-Boschan dating algorithm can now be used to calculate bivariate and multivariate synchronization indices  $\hat{\rho}$  over the entire sample period. Nevertheless, before we proceed with estimating the full-sample multivariate synchronization indices, we would like to test the null hypothesis that all bivariate correlations are equal to 1. Harding and Pagan (2006b) propose two such test, which they call strong positive perfect synchronization (SPPS i) and SPPS ii). The first one tests for equality of the means of the binary series, which constitutes a necessary condition for perfect correlation. If SPPS i) is not rejected they propose to proceed with the SPPS ii) test, which has the null hypothesis that  $\rho_{ij} = 1$  for all  $i$  and  $j$ . For details of the testing procedure we refer to Harding and Pagan (2006b). In Table 3.1 in the last two columns we report the two test statistics for country combinations. We can never reject the SPPS i) but always reject the SPPS ii) test at a 1% significance level. Rejecting the null hypothesis of perfect correlation among all country combinations we can proceed with the estimation of multivariate synchronization indices being strictly smaller than one. These are also reported in Table 3.1.

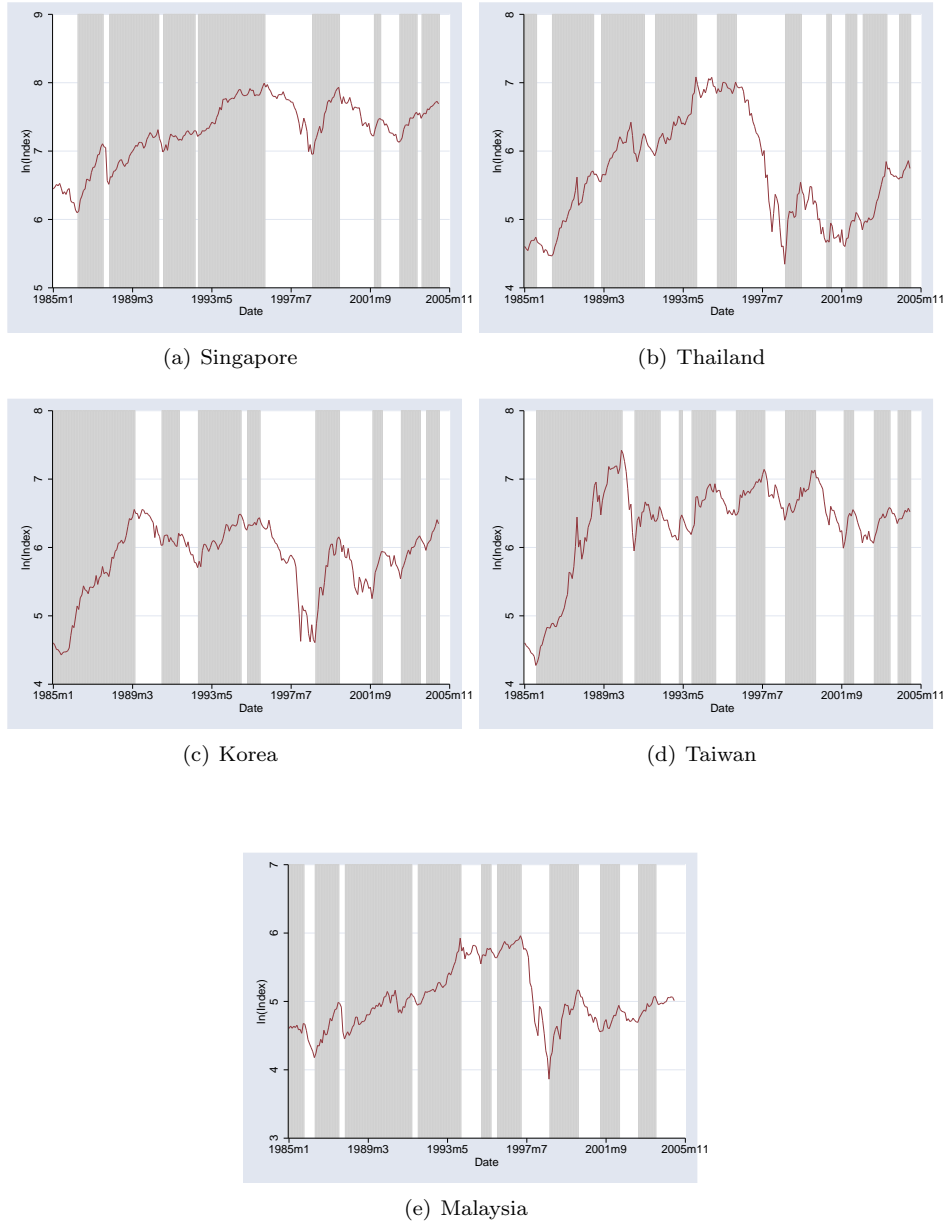
The GMM estimator of the full sample bivariate and multivariate synchronization index  $\hat{\rho}_0$  is calculated using ((3.4)). As noted earlier, the GMM estimator simplifies to the Pearson correlation for bivariate cycle pairs whereas the

<sup>25</sup>The exact dates of the estimated peaks and troughs for each country are not reported in a separate table but are available upon request.

<sup>26</sup>For example, Korean and Thai stock markets already were in a bear phase prior to the Asian crisis for a considerable amount of time (three and one and a half years, respectively). See Edwards et al. (2003) and our own calculations.



Figure 3.6: Monthly Asian stock index prices: Bulls and bears classification



**Note:** Shaded areas correspond with bullish phases.

Table 3.1: Full sample cyclical stock market synchronization in South East Asia: Estimates and homogeneity tests

country sets	$\hat{\rho}$	$[\hat{\rho}_-, \hat{\rho}_+]$	$W(\hat{\rho})$	$CV_W$	SPPS i)	SPPS ii)
Panel A: n=2						
<i>SING; THAI</i>	0.43	[0.18, 0.68]	-	-	0.44	25.99
<i>SING; KOR</i>	0.47	[0.15, 0.79]	-	-	0.59	18.20
<i>SING; TAI</i>	0.34	[0.03, 0.66]	-	-	0.05	27.22
<i>SING; MAL</i>	0.33	[0.02, 0.65]	-	-	0.05	27.07
<i>THAI; KOR</i>	0.31	[0.03, 0.59]	-	-	0.01	34.68
<i>THAI; TAI</i>	-0.04	[-0.36, 0.27]	-	-	0.08	65.50
<i>THAI; MAL</i>	0.43	[0.16, 0.69]	-	-	0.83	26.48
<i>KOR; TAI</i>	0.40	[0.33, 1]	-	-	0.24	24.77
<i>KOR; MAL</i>	0.26	[0.23, 1]	-	-	0.71	33.85
<i>TAI; MAL</i>	0.22	[-0.10, 0.54]	-	-	0.16	36.89
Panel B: n=3						
<i>SING; THAI; KOR</i>	0.42	[0.15, 0.69]	2.81	14.14	0.78	78.88
<i>SING; THAI; TAI</i>	-	-	26.51	18.92	0.47	118.17
<i>SING; THAI; MAL</i>	0.42	[0.12, 0.72]	0.81	18.71	0.94	79.54
<i>SING; KOR; TAI</i>	0.41	[0.10, 0.73]	0.68	14.25	0.62	70.19
<i>SING; KOR; MAL</i>	0.38	[0.03, 0.74]	2.79	18.40	0.86	79.12
<i>SING; TAI; MAL</i>	0.31	[-0.07, 0.69]	1.39	18.78	0.16	91.18
<i>THAI; KOR; TAI</i>	-	-	24.58	15.44	0.24	124.95
<i>THAI; KOR; MAL</i>	0.37	[0.07, 0.66]	1.33	18.34	1.05	95.01
<i>THAI; TAI; MAL</i>	0.32	[0.14, 0.51]	12.99	19.87	0.91	128.87
<i>KOR; TAI; MAL</i>	0.33	[-0.01, 0.67]	2.31	18.18	0.72	95.51
Panel C: n=4						
<i>SING; THAI; KOR; TAI</i>	0.61	[0.52, 0.70]	31.96	36.79	0.79	196.36
<i>SING; THAI; KOR; MAL</i>	0.42	[0.09, 0.75]	6.29	35.33	1.47	166.28
<i>SING; THAI; TAI; MAL</i>	0.53	[0.39, 0.67]	30.58	38.44	1.27	209.15
<i>SING; KOR; TAI; MAL</i>	0.39	[0.00, 0.79]	5.35	32.23	0.86	167.99
<i>THAI; KOR; TAI; MAL</i>	0.50	[0.32, 0.69]	26.20	42.73	1.05	222.17
Panel D: n=5						
ALL	0.60	[0.36, 0.84]	36.78	81.90	1.55	320.65

**Note:** The estimated common synchronization index is denoted by  $\hat{\rho}$ . The closed interval  $[\hat{\rho}_-, \hat{\rho}_+]$  is the corresponding 95% confidence interval for  $\hat{\rho}$ .  $W(\hat{\rho})$  tests for the  $SMS(\hat{\rho})$  hypothesis (test of “multivariate synchronization” or “homogeneity”).  $CV_W$  stands for the 95% critical value of the bootstrap version of the test. The bootstrap is performed on the binary series in blocks. In the bivariate case  $n = 2$ , tests for multivariate synchronization are meaningless because there is only one bivariate correlation. SPPS i) and SPPS ii) test the null of equality of means and equality of all bivariate cross correlations equal to 1, respectively. All test statistics for SPPS i) are insignificant. Whereas all test statistics for SPPS ii) are significant at a 1% level. For details about the SPPS test we refer to Harding and Pagan (2006b).

multivariate synchronization indices stand for the restricted value of the bivariate Pearson correlation in higher dimensions, provided that the null hypothesis  $\text{SMS}(\rho_0)$  is not rejected, i.e.,  $W(\hat{\rho}_0) \leq CV_W(95\%)$ . The value for the test statistic and the critical value  $CV_W$  are also reported in the table.<sup>27</sup> The critical value is determined using a block bootstrap. Finally, the reported closed intervals  $[\rho_-, \rho_+]$  contain all values of  $\rho_0$  that lead to nonrejection of the null hypothesis of  $\text{SMS}(\rho_0)$ . In order to better grasp the results in Table 3.1, consider, for example, the index of multivariate synchronization  $\hat{\rho} = 0.42$  for the triplet  $[SING, THAI, KOR]$ . The bivariate correlations for this triplet are of the same order of magnitude indeed. Thus, it should not be surprising that the null hypothesis  $\rho[SING, THAI] = \rho[SING, KOR] = \rho[KOR, THAI]$  cannot be rejected over the interval  $[0.15, 0.69]$  which contains all three bivariate correlations. The nonrejection justifies the “restricted” 0.41 estimate and is obtained by minimizing the W-test over this interval.

If one compares the magnitude of synchronization indices in Table 3.1, the bivariate comovements seem to differ quite a lot at first sight, ranging from -0.04 (THAI, TAI) to 0.659 (all markets). However, the null hypothesis of a common stock market cycle ( $\text{SMS}(\rho_0)$ ) can nearly never be rejected, which justifies the reported multivariate synchronization estimates in the lower panels of Table 3.1. Those cases for which stock market cycles do not exhibit a common (homogeneous) correlation are denoted by “rej”. Notice also that the polar cases  $\rho = 0$  and 1 fall in the nonrejection intervals  $[\rho_-, \rho_+]$  for only seven out of 26 cases. This provides additional justification to allow for an “intermediate” estimation and testing procedure for “imperfect” multivariate synchronization.

We previously argued that cycle correlations - like the ones in Table 3.1 - are the more relevant risk diversification indicators, provided investors base their buying and selling decision on how stock cycles evolve (i.e., if investors look at cycle turning points to time their buying and selling decisions). Moreover, the conventional correlations based on return pairs are found to be quite different for the majority of considered stock market pairs (not reported in the table). In other words, return correlations can provide misleading information about the potential for risk diversification when investors’ time horizon (and thus their portfolio rebalancing) coincides with the stock market cycle.

In Chapter 4 we will see that one should be very careful in interpreting the full sample correlations in Table 3.1 because they hide different subsample values. We therefore also calculated these corresponding values for the subsample starting in 1997, which at this point is an exogenous split of the sample. The subsample

---

<sup>27</sup>Because it only makes sense to test for multivariate synchronization when  $n > 2$ , the columns for the W-test and the critical values are left empty in the bivariate panel.

synchronization indices  $\hat{\rho}$  are reported in the final Table 3.2. Table 3.2 also contains subsample testing outcomes for the null hypothesis of imperfect multivariate synchronization using (3.4). If this subsample application of the W-test does not lead to rejections, we are allowed to calculate multivariate synchronization indices for the subsamples. We do not report subsample results for those markets which, according to the methods in Chapter 4, do not exhibit breaks.

The most striking table feature is that the synchronization index in the second subsample is nearly always much bigger than its pre-1997 counterpart. Moreover, the pre-1997 nonrejection intervals are lower and do not overlap with the post-1997 nonrejection intervals. The confidence intervals actually suggest that pre-1997 synchronization was often insignificantly different from zero whereas post-1997 synchronization was close to perfect ( $\rho = 1$ ) for more than half of the cases. As expected, the full sample synchronization index lies somewhere in between the subsample values and can be considered as a rough average of the true subsample values. Turning to the outcomes of the subsample homogeneity test, we find only a limited number of rejections for the first subsample; but deviations from homogeneity all disappear in the second subsample. Thus, the multivariate synchronization index increases over time and heterogeneity - present in the first subsample on a limited scale - almost completely disappears after 1997. In Chapter 4 we dwell more on possible interpretations and implications of this.

Table 3.2: Subsample cyclical stock market synchronization in South East Asia: Estimates and homogeneity tests

country sets	subsample 1			subsample 2		
	$\hat{\rho}_1$	$[\hat{\rho}_-, \hat{\rho}_+]$	$W(\hat{\rho}_1)$	$CV_W$	$\hat{\rho}_2$	$W(\hat{\rho}_2)$
Panel A: n=2						
<i>SING; THAI</i>						
<i>SING; KOR</i>	-0.26	$[-0.55, 0.03]$		no break	0.76	$[0.61, 0.92]$
<i>SING; TAI</i>	-0.01	$[-0.46, 0.44]$			0.74	$[0.48, 1]$
<i>SING; MAL</i>						
<i>THAI; KOR</i>	-0.21	$[-0.48, 0.08]$		no break	0.54	$[0.26, 0.82]$
<i>THAI; TAI</i>	-0.46	$[-0.53, -0.39]$			0.45	$[0.06, 0.85]$
<i>THAI; MAL</i>						
<i>KOR; TAI</i>	0.04	$[-0.39, 0.46]$		no break	0.87	$[0.73, 1]$
<i>KOR; MAL</i>	-0.16	$[-0.41, 0.08]$			0.62	$[0.23, 1]$
<i>TAI; MAL</i>	-0.21	$[-0.49, 0.06]$			0.65	$[0.24, 1]$
Panel B: n=3						
<i>SING; THAI; KOR</i>	-0.14	$[-0.47, 0.18]$	2.58	41.49	0.73	$[0.51, 0.95]$
<i>SING; THAI; TAI</i>	rej.		116.84	27.17	0.70	$[0.32, 1]$
<i>SING; THAI; MAL</i>				no break		
<i>SING; KOR; TAI</i>	0.08	$[-0.30, 0.46]$	0.74	31.38	0.89	$[0.66, 1]$
<i>SING; KOR; MAL</i>	-0.01	$[-0.14, 0.13]$	10.56	29.50	0.84	$[0.69, 1]$
<i>SING; TAI; MAL</i>	-0.06	$[-0.30, 0.16]$	3.91	24.95	0.96	$[0.80, 1]$
<i>THAI; KOR; TAI</i>	rej.		58.11	29.82	0.85	$[0.65, 1]$
<i>THAI; KOR; MAL</i>	0.03	$[-0.05, 0.12]$	38.09	40.49	0.61	$[0.38, 0.83]$
<i>THAI; TAI; MAL</i>	rej.		160.27	25.73	0.77	$[0.49, 1]$
<i>KOR; TAI; MAL</i>	-0.09	$[-0.38, 0.19]$	2.40	25.87	0.85	$[0.58, 1]$
Panel C: n=4						
<i>SING; THAI; KOR; TAI</i>	-		-	-	-	-
<i>SING; THAI; KOR; MAL</i>	0.11	$[-0.07, 0.29]$	85.61	131.92	0.86	$[0.75, 0.97]$
<i>SING; THAI; TAI; MAL</i>	rej.		235.99	78.43	0.92	$[0.66, 1]$
<i>SING; KOR; TAI; MAL</i>	0.02	$[-0.25, 0.3]$	12.56	77.35	0.94	$[0.79, 1]$
<i>THAI; KOR; TAI; MAL</i>	rej.		336.52	76.39	0.85	$[0.5, 1]$

**Note:** Variables indexed by 1 (resp. 2) refer to the pre- (resp. post) break period.  $\hat{\rho}$  represents estimates of the synchronization index and  $[\rho_-, \rho_+]$  is the corresponding confidence interval at 95%.  $W(\hat{\rho})$  is the test statistic of the  $SM_S(\hat{\rho})$  hypothesis and  $CV_W$  is the 95% critical value of the bootstrap version of the test. The bootstrap is performed on the binary series in blocks of length 25. Those cases in which the stability test could not be adequately performed due to the near-singular character of the variance-covariance matrix are denoted by (-). This is for the sample with all countries  $n = 5$  the case and the reason why we refrain from showing those results in a separate panel. In the bivariate case  $n = 2$ , tests for multivariate synchronization are meaningless because there is only one bivariate correlation.

## 3.6 Conclusions

In this chapter we proposed a Generalized Method of Moments (GMM) framework in the spirit of Harding and Pagan (2006b) in order to measure the degree of synchronization between stock market “bulls” and “bears”. We argued that an assessment of cycle duration and cycle comovement is potentially relevant for investors that base their investment decisions on the turning points of the stock market cycle. Moreover, policy makers and regulators might be interested to know the magnitude of stock market synchronization because of the potentially destabilizing effects of stock market bulls and bears on the real economy.

Prior to calculating a measure of cyclical synchronization, we classified stock prices into bull and bear periods using the Bry and Boschan (1971) dating algorithm. We subsequently extended the Harding and Pagan (2006b) framework in several directions.

We allowed for a value of the common synchronization index between -1 and 1 whereas Harding and Pagan (2006b) only tested against the benchmark cases of complete perfect synchronization or nonsynchronization (so, they did not really estimate a common synchronization index but restricted it to either 0, -1 or +1 prior to performing the test for a common cycle). However, in practice, business cycles as well as financial cycles are neither perfectly synchronized nor completely independent, which implies that our testing framework is closer to reality than Harding and Pagan’s. Moreover, our approach also produces an estimate of the multivariate synchronization index  $\rho_0$  ( $-1 < \rho_0 < 1$ ).

Before putting the test at work in an empirical application, we performed a Monte Carlo experiment for a variety of representative Data Generating processes (DGPs) in order to evaluate the (small sample) size and power properties of the procedure. We found that the asymptotic version of the SMS( $\rho_0$ ) test potentially suffers from overrejection under the null hypothesis. We illustrated that the size distortion becomes particularly severe when the number of time series (countries) considered is large relative to the length of the time series. As a remedy to this problem, we proposed a block bootstrap procedure to determine the small sample critical values of the GMM test. The bootstrap is purely nonparametric in that we do not make any parametric assumptions about the DGP of the financial cycle. Moreover, the bootstrap is performed in blocks such as to preserve the persistence of the cycles and their co-cyclicity in the bootstrapped samples. We also proof the consistency of such a bootstrap procedure.

The bootstrap was shown to reduce the size distortion to a satisfactory level and to produce accurate critical values provided the time series length  $T$  is sufficiently long relative to the number of countries  $N$ . In the small sample power study of the bootstrap-based test, we made use of the small sample critical values such

as to obtain size-adjusted power values. We found that the test for imperfect multivariate synchronization of order  $\rho_0$  ( $\text{SMS}(\rho_0)$ ) already exhibits satisfactory small sample power against close alternatives for  $\rho_0$ .

We applied the financial cycle synchronization framework to five East-Asian economies. First, we were able to reject the Harding and Pagan (2006b) test towards Strong Perfect Positive Synchronization (SPPS) in a majority of cases. Next, we tested the weaker null hypothesis of Strong Multivariate synchronization of order  $\rho_0$  ( $\text{SMS}(\rho_0)$ ). Using bootstrapped critical values we found that a common multivariate synchronization index is justified for most of the considered country subsets. This may be interpreted as evidence that East-Asian financial markets are well integrated. Finally, we also investigated the robustness of our results over time, i.e., are the synchronization results time-varying? First best would have been to implement stability testing procedures but we leave this for Chapter 4. We therefore limited ourselves to a comparison of subsample results. We find that post-1997 synchronization is larger than pre-1997 and full sample multivariate synchronization.

---

## ON MEASURING SYNCHRONIZATION OF BULLS AND BEARS: THE CASE OF EAST ASIA

This chapter<sup>1</sup> extends the ideas of Chapter 3 and proposes a test for endogenously determining structural change in the bivariate and multivariate synchronization indices. Upon applying the technique to five Asian stock markets we find a significant increase in the cross-country comovements of Asian bullish and bearish periods in 1997. A power study of the stability test suggests that the detected increase in comovement is more of a sudden nature (i.e. contagion or “Asian Flu”) instead of gradual (i.e. financial integration). It is furthermore argued that stock market cycles and their propensity toward (increased) synchronization contain useful information for investors, policy makers as well as financial regulators.

### 4.1 Introduction

In the past couple of years, financial economists have extensively documented the empirical features of stock market returns such as clusters of volatility and heavy tails, see e.g. Embrechts et al. (1997). That stock prices typically exhibit periods of persistent rises or falls, i.e., so-called “bulls” and “bears”, has been recognized by the financial practitioners for a long time but has attracted much less attention from the academic community.<sup>2</sup> Accordingly, the potential for stock

---

<sup>1</sup>This chapter is based on the paper Candelon, Piplack, and Straetmans (2008b) in the *Journal of Banking and Finance*.

<sup>2</sup>Traditionally, there are two perspectives on stock market bulls and bears. First, they may be induced by irrational “animal spirit” (or sentiment that is unrelated to any rational expectations of future fundamental values), see e.g. Summers (1986), Shiller (2000) or Anderson et al. (2003). These papers argue that prices can sometimes display seemingly persistent deviations from their long-run equilibrium values. Another view states that, although market sentiment can drive prices away from fundamentals in the short run, proportional differences between market prices



markets to be simultaneously bullish or bearish across geographical borders has also stayed underexposed. The main objective of this chapter will therefore be to provide a framework for measuring synchronization between stock market cycles and examine how it has evolved over time.<sup>3</sup>

Measuring stock market cycles and their cross-border synchronization is of potential interest for both investors and policy makers. First, it is common sense that investors rebalance their portfolios by purchasing “cheap” stocks during bearish periods and selling “expensive” stocks when stock markets are bullish. The question arises, however, how to optimally time this portfolio rebalancing. Technical analysts typically time their buying and selling decisions by means of graphs. A more thorough statistical analysis of bulls and bears can help investors to further improve the timing of their investment decisions. The above discussion suggests that the duration of a stock cycle constitutes the natural time horizon for a “single-cycle” or “short-term” investor (or, alternatively, constitutes the natural time horizon for portfolio rebalancing in case of a “multi-cycle” or “long-term” investor). Thus, in order to assess the potential for risk diversification across stock market cycles, it seems natural to consider correlations over the duration of a typical stock market cycle and not on, say, a daily or monthly basis. The latter correlation measures might offer a misleading view on the potential for risk diversification if investors base their rebalancing decisions on stock market cycles.

Also, persistent swings in stock market prices and the potentially destabilizing effects on the real economy raise the issue of how monetary authorities should respond. Indeed, bullish stock markets can induce large amounts of loan collateral - especially in less developed banking systems with poor regulatory frameworks - which then increase demand and goods price inflation. Moreover, when the stock market bulls turn into bears, this can result in widespread liquidity problems and a “credit crunch” in the financial system. Thus, monitoring the impact of stock market swings is also of potential interest to regulatory bodies caring about systemic risk and overall financial stability. Finally, if stock cycles have become more synchronized over time, the potential for financial system instability to spill over to other countries has also increased which suggests that a coordinated effort of policymakers and regulatory bodies is necessary.

---

and fundamentals are kept within bounds and stock markets exhibit a long-run relation between prices and fundamentals, see e.g. DeLong (1992), Siegel (1998) or Coakley and Fuertes (2006). Our research does not fit in either of these two strands of literature as we do not aim to disentangle the causes of stock market bulls and bears.

<sup>3</sup>Few empirical studies identified and investigated univariate features of stock market cycles, see e.g. Edwards et al. (2003), Gómez Biscarri and Pérez de Gracia (2003), Pagan and Sossounov (2003) or Lunde and Timmermann (2004). An even smaller set of papers looked into whether stock market cycles comove, see e.g. Gómez Biscarri and Pérez de Gracia (2003), Edwards et al. (2003) and Harding and Pagan (2006b).

This chapter makes several contributions to the stock market bulls and bears cum synchronization literature.<sup>4</sup> More specifically, we extend the Generalized Method of Moments (GMM) approach to measuring business cycle synchronization due to Harding and Pagan (2006b) toward estimating and testing for (bivariate, multivariate) cyclical stock market synchronization. First, we allow for a value of the “common” synchronization index between -1 and 1, whereas Harding and Pagan only consider tests for the benchmark cases of perfect synchronization or nonsynchronization. Second, the estimation and testing procedure for multivariate synchronization is complemented with an endogenous stability test for detecting time variation in cyclical stock market synchronization. Third, our stability testing procedure can be seen as extending a scant preceding literature on structural change in cyclical stock market synchronization see e.g. Edwards et al. (2003). The latter papers tried to investigate the stability of bivariate concordance indices by means of rolling regressions.

Emerging markets are the more obvious candidates for detecting changes in cyclical stock market synchronization due to the rapid transformation of their financial systems and the recurrent financial crises, see e.g. Bekaert and Harvey (2000). We therefore use Asian stock market data in our empirical application. Furthermore, Monte Carlo simulation reveals that the stability test is able to detect contagion-like processes but not gradual changes. In that sense, our paper complements a recent literature on Asian contagion identification, see e.g. Forbes and Rigobon (2002) or Dungey et al. (2006).

The remaining sections are organized as follows. Section 4.2 discusses the statistical estimation and testing methodology that will be implemented. Section 4.3 evaluates the small sample properties of the stability test (size and power) and proposes a bootstrap method for determining the size-corrected small sample critical values. Empirical estimation and testing results are reported in Section 4.4. Concluding remarks are contained in Section 4.5.

---

<sup>4</sup>Preceding literature on stock market synchronization is vast and difficult to summarize. Early research on stock market linkages mainly documented cross border return correlations (see e.g. Roll (1988)). This correlation analysis was refined later on, either by implementing multivariate stochastic volatility (SV) models or cointegrated Vector Auto Regressions (VAR). Representative articles of the former “school” include King and Wadhwani (1990); Mallaris and Urrutia (1992); Lin et al. (1994); Susmel and Engle (1994). These ARCH-type models were used, *inter alia*, to investigate the direction of international spillovers as well as to identify differences in market comovements in periods of market turbulence and market quiescence. Baillie and Bollerslev (1989), however, argued that the modeling of returns can result in a loss of information on possible common trends when prices are cointegrated. Representative articles of the cointegrated VAR literature are Kasa (1992) and Click and Plummer (2005).

## 4.2 Methodology

The concept of imperfect multivariate equity market synchronization is introduced in Section 4.2.1. Section 4.2.2 proposes a GMM-based testing procedure for detecting structural change in the multivariate synchronization measure.

### 4.2.1 Strong multivariate synchronization of degree $\rho_0$

Harding and Pagan (2006b) introduced the concept of synchronization for measuring business cycle comovements. We will brief their GMM framework together with the extensions of Candelon et al. (2008a).

Prior to calculating cyclical correlations we need to identify what stock market “bulls” and “bears” are. The financial press nowadays usually focuses on increases (declines) of markets being greater (less) than either 20% or 25%, see Pagan and Sossounov (2003). In the academic literature, there is no consensus on what bulls and bears actually mean. One definition describes bull or bear markets as “periods of generally increasing/decreasing market prices”, see Chauvet and Potter (2000, p.90, fn.6). The former definition, by focusing on extreme movements, would be analogous to “booms” and “busts” in the real economy whereas the latter definition seems closer to reflecting business cycle contractions and expansions. We use the latter definition of bulls and bears that focuses on how stock prices evolve between local peaks and troughs. This approach is in line with the business cycle literature going back to Burns and Mitchell (1946). The definition essentially implies that a bullish stock market turned bearish if prices have declined for a substantial period since their previous (local) peak. Such a definition does not exclude sequences of price falls (rises) during a bull (bear) phase, but there are restrictions on the extent to which these sequences of price reversals can occur and yet still be considered part of any given bull or bear phase.

Let  $p_{it}$  denote the log stock price for country  $i$  at time  $t$  ( $i = 1, \dots, n; t = 1, \dots, T$ ). Bull and bear periods are determined using the marginal transform  $\varphi(\cdot)$  such that  $\varphi(p_{it}) = S_{it}$  ( $\forall i$ ) where  $S_{it}$  is 0 or 1 in case of bear or bull period, respectively. There are two main methodological strands in the literature to select  $\varphi(\cdot)$ . First, Hamilton (1989) imposes a two regime Markov-switching model on  $\varphi(\cdot)$  that allows for booms and busts.<sup>5</sup> We prefer a second, nonparametric approach which can be motivated by the complex behavior of financial time series (clusters of volatility, long memory etc.) and the resulting risk of model misspecification. A key feature of nonparametric filters is the location of turning points that correspond to local maxima and minima of the series. Loosely speaking, a peak/trough

---

<sup>5</sup>Applications of this approach include Hamilton and Lin (1996) and Maheu and McCurdy (2000).

in the stock price series occurs when  $p_{it}$  reaches a local maximum (minimum) in a window of six months width.<sup>6</sup> For a peak to occur at  $t$ , this implies:

$$p_{i,t-6} \cdots p_{i,t-1} < p_{i,t} > p_{i,t+1} \cdots p_{i,t+6}.$$

Likewise there will be a trough at  $t$  if

$$p_{i,t-6} \cdots p_{i,t-1} > p_{i,t} < p_{i,t+1} \cdots p_{i,t+6}.$$

The above dating algorithm constitutes the core of a more complicated method proposed by Bry and Boschan (1971) in order to date business cycle phases at the quarterly frequency. This approach has since also been used to determine stock market cycles, see e.g. Kaminski and Schmukler (2008) or Pagan and Sossounov (2003). Once turning points are determined, bull and bear periods can be identified as periods between troughs (peaks) and peaks (troughs), respectively.<sup>7</sup>

Harding and Pagan (2006b) proposed GMM-based multivariate procedures for testing the null hypothesis that business cycles are either perfectly synchronized or not synchronized at all. Candelon et al. (2008a) argued that these two borderline null hypotheses might be a bit unrealistic and proposed a more general framework to test for strong (but imperfect) multivariate synchronization of degree  $\rho_0$  ( $-1 < \rho_0 < 1$ ), or  $\text{SMS}(\rho_0)$  in shorthand notation.<sup>8</sup> Their procedure starts from the following  $n(n+1)/2$  moment conditions under the null hypothesis  $\text{SMS}(\rho_0)$ :

$$E[h_t(\theta_0, S_t)] = 0 \quad (4.1)$$

with

$$h_t(\theta_0, S_t) = \begin{bmatrix} S_{1t} - \mu_1 \\ \vdots \\ S_{nt} - \mu_n \\ \frac{(S_{1t} - \mu_1)'(S_{2t} - \mu_2)}{\sqrt{\mu_1(1-\mu_1)\mu_2(1-\mu_2)}} - \rho_0 \\ \vdots \\ \frac{(S_{(n-1)t} - \mu_{n-1})'(S_{nt} - \mu_n)}{\sqrt{\mu_{n-1}(1-\mu_{n-1})\mu_n(1-\mu_n)}} - \rho_0 \end{bmatrix} \quad (4.2)$$

<sup>6</sup>Robustness checks for different window widths are not included for sake of space considerations. Turning point locations only change marginally.

<sup>7</sup>Throughout the rest of the paper we allow for two regimes, but generalizing the framework is straightforward.

<sup>8</sup>The borderline case of perfect multivariate synchronization ( $\rho_0 = \pm 1$ ) is excluded from the analysis because it is unlikely to be observed and the asymptotic distribution in this limiting case turns out to be a weighted average of  $\chi^2$  distributions with the weights to be determined by simulation, see Gouriéroux et al. (1982).

and with  $\theta'_0 = [\mu_1, \dots, \mu_n, \rho_0, \dots, \rho_0]$ , the restricted vector under the  $\text{SMS}(\rho_0)$  case. The expectations operator in (4.1) is defined over the time dimension. The first subset of  $n$  moment conditions in (4.1) and (4.2) defines the population means of the cycle dummies and reflects the likelihoods of the stock markets to be in the bullish phase. The remaining  $n(n-1)/2$  moment conditions express equality of all cyclical correlations to some common value  $\rho_0$ . If the above moment conditions hold, stock markets, albeit imperfectly synchronized, do exhibit a “common” or “homogeneous” synchronization index  $\rho_0$ .

In Chapter 3 we propose to test  $\text{SMS}(\rho_0)$  via Hansen’s (1982) Wald test statistic:

$$W(\rho_0) = Tg(\theta_0, [S]_{t=1}^T)' \hat{V}^{-1} g(\theta_0, [S]_{t=1}^T), \quad (4.3)$$

which converges to an asymptotic  $\chi^2_{n(n-1)/2}$  distribution under the null hypothesis  $\theta = \theta_0$ , see e.g. Harding and Pagan (2006b, p.70). The covariance matrix  $\hat{V}$  is a heteroskedasticity and autocorrelation consistent (HAC) estimator of the variance-covariance matrix of  $\sqrt{T}g(\theta_0, [S]_{t=1}^T)$ , see e.g. Newey and West (1987). The statistic is defined as a quadratic form in the penalty vector  $g(\theta_0, [S]_{t=1}^T) = \frac{1}{T} \sum_{t=1}^T h_t(\theta_0, S_t)$  which reflects the average deviation from the moment conditions. The stronger the deviations from the moment conditions the more likely a rejection of  $\text{SMS}(\rho_0)$  becomes.

Notice that the implementation of the Wald test (4.3) still hinges upon an (unknown) value for  $\rho_0$ . Here we propose a two-step estimation procedure: first, they determine the closed interval  $[\rho_-, \rho_+]$  for which  $\text{SMS}(\rho_0)$  cannot be rejected at a prespecified nominal size. Next, a natural estimator for  $\rho_0$  follows by minimizing the test statistic in (4.3) over  $[\rho_-, \rho_+]$ :

$$\hat{\rho}_0 = \text{Argmin}_{\rho \in [\rho_-, \rho_+]} \left\{ Tg(\theta_0, [S]_{t=1}^T)' \hat{V}^{-1} g(\theta_0, [S]_{t=1}^T) \right\}. \quad (4.4)$$

In the next subsection we develop a framework to test for structural change in  $\hat{\rho}_0$ .

## 4.2.2 Structural breaks in multivariate synchronization

To the best of our knowledge, very few studies on co-cyclicity in time series have yet considered the possibility of time variation in co-cyclicity.<sup>9</sup> This may be partly due to the relative complexity of performing stability tests within a GMM framework. The full sample framework of Harding and Pagan (2006b)

<sup>9</sup>A few studies look at temporal stability of so-called “concordance” indices, see e.g. Edwards et al. (2003), or Gómez Biscarri and Pérez de Gracia (2003). However, the synchronization measure has several advantages over the concordance index. For a more in-depth discussion of the pitfalls of concordance indices, see e.g. Harding and Pagan (2006b, p.65).

makes the simplifying assumption that the cross sectional dependence structure of raw stock returns  $(\Delta p_{1t}, \dots, \Delta p_{it}, \dots, \Delta p_{nt})$  and corresponding cycle dummies  $(S_{1t}, \dots, S_{it}, \dots, S_{nt})$  remains stationary over the entire sample period. However, assuming stationarity seems problematic, especially for emerging markets that have been recently characterized by financial liberalization experiments (e.g. gradual abolishment of capital controls), recurrent changes in domestic supervisory rules, exchange rate regimes and severe financial crises, see e.g. Bekaert and Harvey (2000) for details on the reform processes in emerging countries.

To start the breakpoint analysis, we concentrate on the simplifying case of a known breakpoint at time  $b = T_1/T$ , i.e., a priori selected by the researcher. This implies that the full sample moment conditions in (4.1)-(4.2) become:

$$\begin{cases} ES_{it} = \mu_i & i = 1, \dots, n \\ E \left[ \frac{(S_{it} - \mu_i)'(S_{jt} - \mu_j)}{\sqrt{\mu_i(1-\mu_i)\mu_j(1-\mu_j)}} - \rho_1^{ij} \right] = 0 & i = 1, \dots, n, j \neq i, t = 1, \dots, T_1 \\ E \left[ \frac{(S_{it} - \mu_i)'(S_{jt} - \mu_j)}{\sqrt{\mu_i(1-\mu_i)\mu_j(1-\mu_j)}} - \rho_2^{ij} \right] = 0 & i = 1, \dots, n, j \neq i, t = T_1 + 1, \dots, T. \end{cases} \quad (4.5)$$

The latter set of moment conditions further differs from (4.1)-(4.2) in that (4.5) allows for correlation differences across stock market pairs  $ij$ .<sup>10</sup> This allows us to perform a restricted and an unrestricted version of the stability test. The null hypothesis of temporal stability in bivariate synchronization corresponds to the case where  $\rho_1^{ij} = \rho_2^{ij}$  ( $i = 1, \dots, n; j \neq i$ ). We call this null hypothesis the “weak” or “unrestricted” stability hypothesis as it does not require cross sectional equality of the bivariate synchronization index in each subsample (the existence of a multivariate synchronization index). On the other hand, the “strong” or “restricted” stability hypothesis,  $\rho_1 = \rho_2$ , considers structural change in the multivariate synchronization index; this version of the test presupposes the cross sectional equality of bivariate correlations in each of the two subsamples (“homogeneous” bivariate synchronization). In order to disentangle changes in multivariate synchronization from time varying deviations in homogeneity, we will test the unrestricted version of the stability hypothesis in combination with subsample tests for multivariate synchronization (see previous section for a discussion of the latter test). If homogeneity holds in the subsamples defined by the break, both the weak and strong stability test should render the same break estimates.

<sup>10</sup>More specifically,  $\rho_1^{ij}(\rho_2^{ij})$  stands for the bivariate Pearson correlation of the binary pair  $(S_{it}, S_{jt})$  in the pre- (post-) break period.

To facilitate the presentation of the stability test, it is worthwhile to introduce some additional notation. Let  $\hat{\theta}'_k(b) = [\hat{\mu}_{k1}, \dots, \hat{\mu}_{kn}, \hat{\rho}_k^{12}, \dots, \hat{\rho}_k^{n(n-1)}]$  be the (unrestricted) GMM estimator of  $\theta'$  based on subsamples  $k = 1, 2$  with  $b = T_1/T$ . The natural statistic for testing the unrestricted or restricted stability hypothesis is the Wald statistic proposed by Andrews and Fair (1988):

$$\Sigma_T(b) = T[\hat{\theta}_1 - \hat{\theta}_2]' \hat{V}_S(b)^{-1} [\hat{\theta}_1 - \hat{\theta}_2],$$

where

$$\hat{V}_S(b) = (1/b)[\hat{F}_1(b)' \hat{S}_1^{-1}(b) \hat{F}_1(b)]^{-1} + (1/(1-b))[\hat{F}_2(b)' \hat{S}_2^{-1}(b) \hat{F}_2(b)]^{-1},$$

and

$$\hat{F}_1(b) = [T_1]^{-1} \sum_{t=1}^{T_1} \frac{\partial h_t(\hat{\theta}_1(b), S_t)}{\partial \theta'}, \quad \hat{F}_2(b) = [T - T_1]^{-1} \sum_{t=T_1+1}^T \frac{\partial h_t(\hat{\theta}_2(b), S_t)}{\partial \theta'}$$

As in the previous section, the deviations from the (subsample) moment conditions in (4.5) are defined by  $g(\hat{\theta}_1(b), \{S_t\}_{t=1}^{T_1}) = \frac{1}{T_1} \sum_{t=1}^{T_1} h_t(\hat{\theta}_1, S_t)$  and  $g(\hat{\theta}_2(b), \{S_t\}_{t=T_1+1}^T) = \frac{1}{T-T_1} \sum_{t=T_1+1}^T h_t(\hat{\theta}_2, S_t)$ . Andrews and Fair (1988) showed that  $\Sigma_T(b)$  converges to a  $\chi_{n(n-1)/2}^2$  distribution.

We are now ready to consider the design of optimal tests for parameter variation at some unknown breakpoint. In line with Quandt's (1960) pioneering work on endogenous breakpoint determination in linear time series models,  $\Sigma_T(b)$  can be calculated to produce a sequence of statistics indexed by  $b$ . This sequence can be used to construct a single statistic for testing the null hypothesis. Andrews (1993) suggests to select  $b$  such as to maximize  $\Sigma_T(b)$

$$Q_A(b^*) = \sup_{b \in B} \left\{ T[\hat{\theta}_1 - \hat{\theta}_2]' \hat{V}_S(b)^{-1} [\hat{\theta}_1 - \hat{\theta}_2] \right\}. \quad (4.6)$$

At the candidate break date  $b^*$  the constancy hypothesis is most likely to be violated.<sup>11</sup> Asymptotic critical values for  $Q_A$  are provided in the same paper.

Based on Quandt (1960) basic idea, Andrews and Ploberger (1994) proposed two tests (the simple average and exponential tests) that complement  $Q_A$ . However, upon comparing their performance, Hall (2005, pp.183-184) concluded that "no one test dominates the others". We therefore limit our analysis to the Supremum functional. Before putting this stability test to work in an empirical application, we establish the small sample properties of the test which will reveal that

---

<sup>11</sup>In accordance with Andrews (1993), the interval  $B$  is chosen equal to the closed interval  $[0.15T; 0.85T]$  where  $T$  represents the total sample size.

blindly using asymptotic critical values from Andrews (1993) can be problematic when the dimension of the test (large number of markets or countries) grows large.

### 4.3 Monte Carlo experiments

Previous work (in particular Christiano and den Haan (1996)) argues that existing asymptotic theory for GMM estimators may break down in small samples. This problem might be more severe when the number of countries grows large relative to the length of the economic time series. In Chapter 3 we show that the asymptotic version of the test for multivariate cycle synchronization of degree  $\rho_0$  ( $SMS(\rho_0)$ ) in (4.3) is biased towards rejection for simulated processes that are representative for the business cycle and financial literature. They therefore propose to bootstrap the small sample distribution of the test. In this section we perform a similar simulation study of size and power for the stability test (4.6) and for a number of data generating processes representative of financial data.

It is well known that unit root processes for the stock price are able to generate persistent rises or falls in simulated prices, see e.g. Pagan and Sossounov (2003). The unit root process  $p_{it} = p_{it-1} + u_{it}$  will therefore act as a benchmark model.<sup>12</sup> The Monte Carlo investigation evaluates the impact that changes in the distributional assumptions for  $\Delta p_{it} = u_{it}$  have on the size and power of the stability test. First, returns  $\Delta p_{it}$  are drawn from a standard normal df.<sup>13</sup> However, the normal df does not capture the fat tail feature of financial asset returns, see e.g. Mandelbrot (1963) or Embrechts et al. (1997). We therefore also allow for student-t distributed innovations in the unit root process. Finally, in order to capture the clusters of volatility feature of financial returns (see e.g. Bollerslev (1986)), we used Bollerslev's Constant Conditional Correlation (CCC) model as simulation vehicle for the price innovations  $u_{it}$ . The mean equation, variance and covariance boil down to:

$$\begin{aligned}\Delta p_{i,t} &= u_{i,t} = \sqrt{h_{ii,t}}\epsilon_{i,t} \\ h_{ii,t} &= 0.01 + 0.2u_{i,t-1}^2 + 0.79h_{ii,t-1},\end{aligned}$$

and

$$h_{ij,t} = 0.5\sqrt{h_{ii,t}h_{jj,t}},$$

where the parameters driving the conditional volatilities are chosen such that the unconditional variance is equal to one (the same unconditional variance as with

<sup>12</sup>For sake of simplicity we assume a zero drift term. However, our simulation results hardly change in the presence of nonzero drift.

<sup>13</sup>This special case for  $p_{it}$  is also called an i.i.d. Gaussian Random Walk.



the normal and student-t innovations). The standardized residuals  $\epsilon_{i,t}$  will either be standard normally distributed or student-t distributed. A well known drawback of the classic raw return correlation constitutes its sensitivity to increases in subsample volatility. As a result, Forbes and Rigobon (2002) argue that correlations should first be corrected for this volatility bias before one can interpret rises in correlation as providing evidence for either financial contagion or changes in interdependence. The CCC model framework enables us to investigate whether the same bias problem arises for the synchronization index.

Once we generate the data, we apply the Bry and Boschan (1971) dating algorithm to determine the dummy pairs  $(S_{i,t}, S_{j,t})$ . Without loss of generality we limit ourselves to considering the size of the stability test under the null hypothesis of strong multivariate non-synchronization (*SMNS*) (i.e.  $E(u_{i,t}u_{j,t}) = 0$  and  $E(\epsilon_{i,t}\epsilon_{j,t}) = 0$  in the described simulation procedures).<sup>14</sup>

Table 4.1 contains the small sample size of the stability test (4.6) for a varying number of countries. The nominal size is set to 5%. We report the small sample size that corresponds with the asymptotic critical values (“noncorrected”) and with the bootstrapped critical values (“bootstrapped”). The sample size  $T$  is set at 250 which roughly corresponds with the sample size in the empirical application. The upper panel reports simulation outcomes for normally distributed and student-t distributed returns whereas the lower panel is generated by simulating the CCC model. The table further distinguishes between unrestricted and restricted stability tests. In the former case one leaves bivariate correlations unrestricted in the pre-break and post-break vectors  $\theta_1$  and  $\theta_2$  whereas the restricted stability test is only implemented for those simulation replications that are consistent with a common synchronization index over the full sample, i.e., when the null hypothesis  $\text{SMS}(\rho_0)$  is not rejected. Finally, the bootstrap is performed on the binary series and not on the original data. Moreover, and in order to take into account the temporal persistence in the cycle variables, we bootstrap in blocks of length 25.<sup>15</sup>

First and foremost, the table shows that small sample size distortions using the asymptotic critical values (“noncorrected” rows) are present and growing with the number of countries. If one tests for structural stability in  $\theta$  for five countries simultaneously, the asymptotic GMM testing procedure nearly always rejects the

---

<sup>14</sup>Fairly similar size distortions arise when simulating under the null hypothesis  $\text{SMS}(\rho_0)$ , with  $\rho_0 \neq 0$ . The latter simulations are therefore omitted but available upon request.

<sup>15</sup>There is no rule-of-thumb for choosing the optimal block length for a bootstrap in a GMM context. Hall et al. (1995) established optimal block lengths for some stochastic processes but none of them is representative of the GMM framework we are using. Thus, it should not be surprising that we found that bootstraps with block lengths conforming to the Hall et al. criteria only partially eliminate the size distortion. As an alternative strategy, we compared the performance of the block bootstrap for a grid of block lengths and found by trial and error that block sizes of 25 more or less eliminate the size distortion.

(true) null hypothesis of absence of breaks in multivariate synchronization. The size distortions are somewhat less severe in the restricted case but still too big to justify the use of asymptotic critical values. In order to remedy the size distortion we perform a block bootstrap along the lines of Chapter 3. The bootstrapped critical values exceed their asymptotic counterparts and bring down the small sample size in the neighborhood of 5% in the majority of the cases. Further details on the block bootstrap procedure can be found in the appendix to this chapter.

Table 4.1: Small sample size of stability test: Uncorrected and bootstrap-corrected

	Unrestricted stability test				Restricted stability test			
	n=2	n=3	n=4	n=5	n=2	n=3	n=4	n=5
Panel A: identically/independently distributed (IID) returns								
uncorrected ( $\alpha = 1$ )	0.164	0.926	1.000	1.000	0.078	0.293	0.506	0.712
uncorrected ( $\alpha = 2$ )	0.161	0.917	1.000	1.000	0.065	0.238	0.483	0.626
uncorrected ( $\alpha = 3$ )	0.178	0.919	1.000	1.000	0.061	0.312	0.524	0.675
uncorrected ( $\alpha = \infty$ )	0.141	0.885	1.000	1.000	0.068	0.258	0.445	0.664
bootstrapped ( $\alpha = 1$ )	0.057	0.055	0.081	0.084	0.061	0.054	0.049	0.053
bootstrapped ( $\alpha = 2$ )	0.048	0.063	0.084	0.049	0.052	0.056	0.068	0.057
bootstrapped ( $\alpha = 3$ )	0.074	0.054	0.069	0.048	0.047	0.061	0.051	0.054
bootstrapped ( $\alpha = \infty$ )	0.058	0.059	0.051	0.053	0.049	0.051	0.039	0.069
Panel B: CCC model								
uncorrected ( $\alpha = 1$ )	0.177	0.932	1.000	1.000	0.068	0.307	0.496	0.706
uncorrected ( $\alpha = 2$ )	0.148	0.926	1.000	1.000	0.067	0.294	0.521	0.686
uncorrected ( $\alpha = 3$ )	0.156	0.927	1.000	1.000	0.071	0.244	0.516	0.701
uncorrected ( $\alpha = \infty$ )	0.144	0.901	1.000	1.000	0.048	0.207	0.474	0.672
bootstrapped ( $\alpha = 1$ )	0.061	0.063	0.055	0.059	0.044	0.045	0.053	0.049
bootstrapped ( $\alpha = 2$ )	0.062	0.061	0.082	0.048	0.041	0.066	0.049	0.052
bootstrapped ( $\alpha = 3$ )	0.064	0.069	0.062	0.058	0.053	0.043	0.039	0.062
bootstrapped ( $\alpha = \infty$ )	0.061	0.059	0.065	0.051	0.048	0.062	0.063	0.053

**Note:** The nominal size is set to 5%. The table entry  $n$  refers to the number of countries. The sample size  $T$  is set to 250. Simulations are performed under the null hypothesis of a zero synchronization correlation  $\rho$ . The upper panel contains simulation results for identically and independently distributed returns generated with the normal df and the student-t df. The degrees of freedom parameter can be equal to 1, 2 or 3 in the latter case and is infinity for the normal df. The lower panel is based on simulations from a CCC model with either normal or student-t distributed innovations. As in the upper panel, the student-t innovations can have three different degrees of freedom. “Unrestricted” tests leave the bivariate correlations free in the pre-break and post-break vectors  $\theta_1$  and  $\theta_2$  in contrast to the restricted tests. Rejection frequencies are reported using the asymptotic (“uncorrected”) as well as the bootstrapped critical values (“bootstrap”). The bootstrap is performed on the binary series in blocks of length 25. We perform 1,000 Monte Carlo replications in conjunction with a variable number of bootstrap draws in each replication.

The table further reveals that the magnitude of the asymptotic size distortions as well as the performance of the bootstrap algorithm do not seem to be influenced by the choice of stochastic process in the simulation. More specifically, this implies, inter alia, that volatility clusters (Panel B) do not induce extra overrejections of the stability hypothesis as compared to models without conditional volatility changes (Panel A). In other words, the synchronization index and accompanying stability

test do not need Forbes and Rigobon (2002) type corrections for clusters of high volatility.<sup>16</sup>

Using size-corrected critical values obtained via the bootstrap, we can now also evaluate the size-corrected power of the stability test. Our power study assumes different adjustment paths and different adjustment speeds for moving from the pre-break synchronization index  $\rho_1$  to its post-break counterpart  $\rho_2$ . First, a “sudden” jump (alternative hypothesis ( $H_{SJ}$ )) corresponds with an instantaneous increase in the synchronization index from  $\rho_1$  to  $\rho_2$  at time  $T/2$ . Second, a “quick” increase (alternative hypothesis ( $H_{QI}$ )) corresponds to a rise of the synchronization index of  $\frac{4(\rho_2 - \rho_1)}{T}$  per time unit, meaning that  $\rho_2$  is reached at  $3T/4$ . Finally, a “slow” increase (alternative hypothesis ( $H_{SI}$ )) corresponds to a rise of the synchronization index of  $\frac{2(\rho_2 - \rho_1)}{T}$  per time unit, which implies that  $\rho_2$  is only reached at the end of the sample  $T$ . These three adjustment speed scenarios could be linked to the concepts of financial integration and contagion. Although there is not a unifying definition in the literature for either of these concepts, financial integration is usually associated with permanent but gradual changes in comovement measures whereas contagion is more associated with sudden but transitory shocks in the dependence structure of financial returns. A power study will clarify whether the test can distinguish gradual changes from sudden changes in the synchronization index.

Corresponding power results are reported in Table 4.2, where we again chose a nominal size of 5%. The power is only evaluated for the unrestricted version of the stability test.

We find that the power is small for changes in  $\rho$  that are either small or gradual (or both). The power is only of an acceptable magnitude for sudden, big changes in  $\rho$ . Indeed, it is surprising to see that the power rapidly worsens when the adjustment speed is lowered (even for large changes in  $\rho$ ). These results suggest that only sudden jumps can be detected which seems to exclude the financial integration interpretation for the breaks. Thus, we can safely conclude that breaks - if detected - provide evidence against gradual integration processes. However, the stability test cannot be considered as a full fledged contagion vs. integration test because it is unable to distinguish between permanent and transitory shocks.

---

<sup>16</sup>This robustness result can be understood as arising from the heteroskedastic and autocorrelation consistent (HAC) nature of the variance-covariance matrix in the multivariate synchronization test and the stability test.

Table 4.2: Small sample (size-corrected) power of bootstrap stability test

$(\rho_0, \rho_1, \alpha)$	IID returns				CCC			
	n=2	n=3	n=4	n=5	n=2	n=3	n=4	n=5
Panel A: sudden jump ( $H_{SJ}$ )								
(0, 0.275, 1)	0.171	0.164	0.113	0.091	0.121	0.139	0.121	0.121
(0, 0.275, 2)	0.164	0.137	0.093	0.055	0.153	0.158	0.134	0.084
(0, 0.275, 3)	0.160	0.131	0.074	0.048	0.154	0.165	0.091	0.053
(0, 0.275, $\infty$ )	0.154	0.138	0.083	0.048	0.147	0.149	0.088	0.048
(0, 0.7, 1)	0.445	0.449	0.322	0.268	0.325	0.301	0.229	0.219
(0, 0.7, 2)	0.57	0.552	0.431	0.321	0.509	0.464	0.378	0.319
(0, 0.7, 3)	0.563	0.574	0.427	0.358	0.536	0.512	0.419	0.347
(0, 0.7, $\infty$ )	0.613	0.609	0.495	0.384	0.568	0.579	0.461	0.413
Panel B: quick increase ( $H_{QI}$ )								
(0, 0.275, 1)	0.142	0.147	0.094	0.057	0.111	0.124	0.115	0.145
(0, 0.275, 2)	0.132	0.107	0.057	0.041	0.134	0.132	0.088	0.078
(0, 0.275, 3)	0.159	0.117	0.066	0.034	0.145	0.128	0.069	0.038
(0, 0.275, $\infty$ )	0.147	0.108	0.066	0.039	0.142	0.122	0.055	0.039
(0, 0.7, 1)	0.346	0.317	0.203	0.169	0.268	0.318	0.259	0.234
(0, 0.7, 2)	0.436	0.386	0.271	0.224	0.419	0.359	0.256	0.202
(0, 0.7, 3)	0.474	0.426	0.268	0.170	0.444	0.368	0.247	0.171
(0, 0.7, $\infty$ )	0.504	0.448	0.272	0.198	0.463	0.413	0.289	0.209
Panel C: slow increase ( $H_{SI}$ )								
(0, 0.275, 1)	0.112	0.098	0.061	0.041	0.091	0.102	0.101	0.122
(0, 0.275, 2)	0.091	0.086	0.045	0.042	0.098	0.096	0.082	0.042
(0, 0.275, 3)	0.087	0.086	0.049	0.031	0.119	0.077	0.045	0.032
(0, 0.275, $\infty$ )	0.092	0.075	0.057	0.034	0.098	0.083	0.049	0.035
(0, 0.7, 1)	0.212	0.164	0.122	0.087	0.145	0.172	0.171	0.161
(0, 0.7, 2)	0.199	0.165	0.112	0.066	0.209	0.183	0.131	0.077
(0, 0.7, 3)	0.213	0.179	0.125	0.061	0.199	0.201	0.098	0.067
(0, 0.7, $\infty$ )	0.231	0.174	0.102	0.061	0.192	0.184	0.109	0.069

**Note:** The nominal size is 5%.  $n$  refers to the number of countries and the sample size  $T$  is set to 250. All three panels contain power results for i.i.d. and CCC model returns generated with the normal df and the student-t df. The degrees of freedom parameter can be equal to 1, 2 or 3 in the latter case and is infinity for the normal df. The break is at  $T/2$ . Panel A, B, and C describe the results for the alternative hypotheses ( $H_{SJ}$ ), ( $H_{QI}$ ), ( $H_{SI}$ ) corresponding to the “sudden jump”, the “quick increase” and the “slow increase” hypotheses, respectively. The synchronization index takes the value  $\rho_0$  before the break and  $\rho_1$  at time  $T$ . Bootstrap draws are performed on the binary series in blocks of length 25. We performed 1,000 Monte Carlo replications and we used a variable number of bootstrap draws in each replication.

## 4.4 Empirical evidence

Here we apply the multivariate synchronization measure and stability test to Asian stock markets. We estimate full sample bivariate and multivariate synchronization indices for stock market cycles of five East Asian countries. We distinguish full sample and subsample results to identify temporal shifts in synchronization. Emerging (Asian) countries are more likely to show changes in the synchronization of stock market cycles than developed markets because of an ongoing financial liberalization process and recurrent financial turmoil. Full sample and subsample estimation results are complemented with a battery of tests. First, we test whether there is parameter variation in the full sample bivariate and multivariate synchronization indices. Second, we run subsample homogeneity tests in order to determine whether the multivariate synchronization (homogeneity) hypothesis breaks down in subsamples. Nonrejection of the homogeneity hypothesis could be interpreted as evidence for “common” stock market cycles. If the common cycle hypothesis does not hold for all considered stock markets, it might still be the case that it holds for a narrower subset, i.e., an Asian “core”. We therefore also pay attention to that possibility.

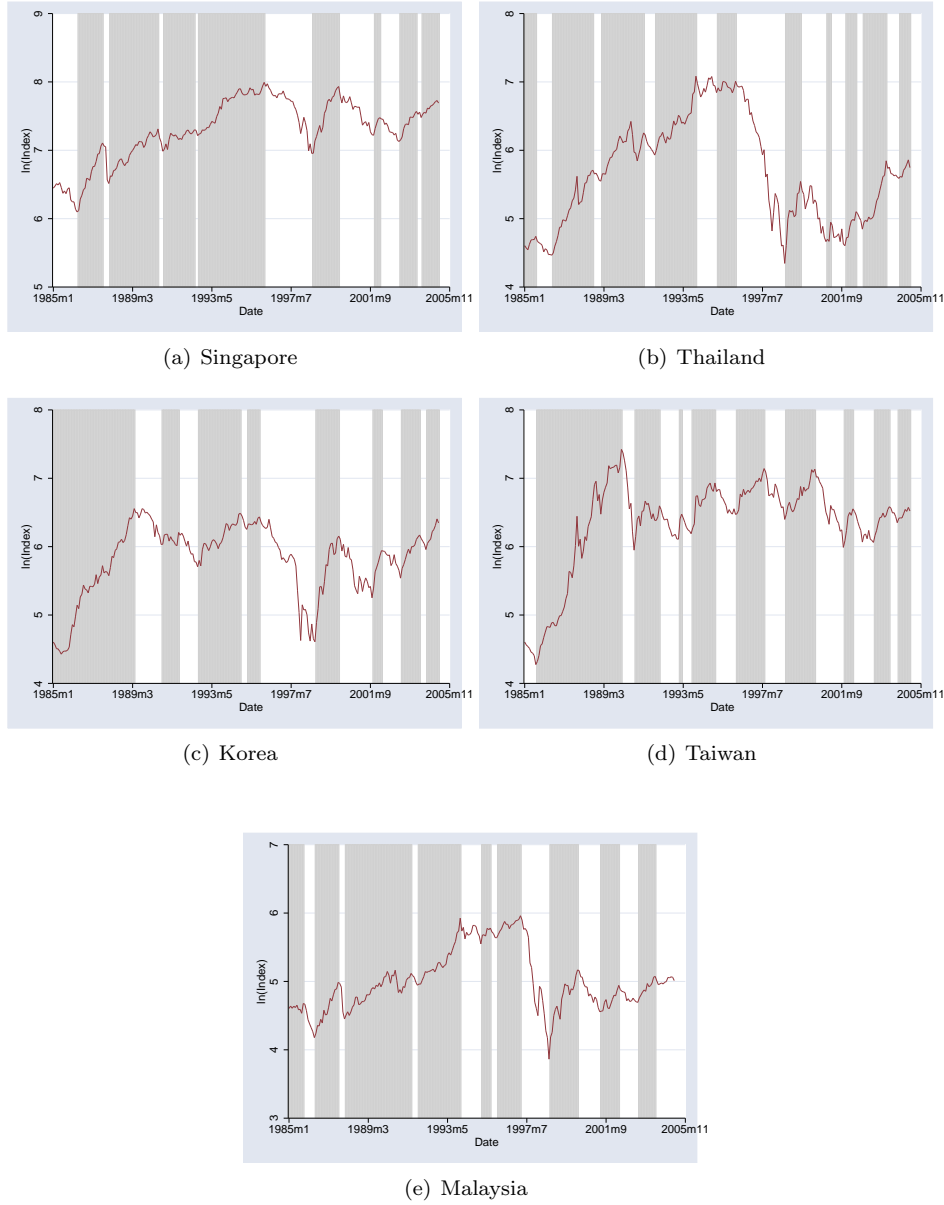
US dollar-denominated and dividend-adjusted monthly stock market indices for Singapore (S), Thailand (Th), South Korea (K), Taiwan (T) and Malaysia (M) were downloaded from the IFS database over the period January 1985 until November 2005 which amounts to 239 monthly observations. Because the prime focus of the paper is a structural change analysis of cyclical stock market comovements, we did not try to maximize the number of countries but selected those Asian countries with the longest possible time series in the database. As we are interested in measuring the comovement of medium-run fluctuations or cycles across stock markets, we first have to identify these “bulls” and “bears”. The cycle dummies are obtained by implementing the Bry and Boschan (1971) dating algorithm over a six month time interval as in Chapter 3. Previous studies seem to suggest that stock market cycle dating and the resulting cyclical comovement measures are relatively robust across different types of (parametric, nonparametric) dating algorithms, see e.g. Candelon et al. (2008a). Figure 4.1 contains the evolution of the (log) stock indices for each of the countries, where the bull periods have been shaded to facilitate visual inspection.<sup>17</sup>

It is striking that the first half of the sample primarily consists of bull periods which illustrates why investors have been talking about “Asian Tigers”. Also, the 1997 Asian crisis and its aftermath are clearly visible in the plots. More specifically, our results replicate the earlier finding that financial crises seem to erupt several

---

<sup>17</sup>The exact dates of the estimated peaks and troughs for each country are not reported in a separate table but are available upon request.

Figure 4.1: Monthly Asian stock index prices: Bulls and bears classification



**Note:** Shaded areas correspond to bullish phases.

months into bear phases (and sometimes very close to the end), see e.g. Edwards et al. (2003).<sup>18</sup> However, and as Edwards et al. (2003, p.936) pointed out, it would be premature to conclude on the basis of this observation that bear markets are leading indicators of financial crises. Somewhat surprisingly, the dotcom bubble burst is also clearly visible despite the relative underrepresentation of technology companies in emerging markets. Last but not least, the figures provide casual evidence for “synchronization” between bull and bear periods across markets. In order to assess the its degree and if it changed over time, we have to resort to the more advanced statistical tools introduced in the previous sections.

The cycle dummies  $S_{i,t}$  ( $i = 1, \dots, n; t = 1, \dots, T$ ) that result from applying the Bry-Boschan dating algorithm can now be used to calculate bivariate and multivariate synchronization indices  $\hat{\rho}$  over the entire sample period. These are reported in Table 4.3.

The GMM estimator of the full sample bivariate and multivariate synchronization index  $\hat{\rho}_0$  is calculated using (4.4). As noted earlier, the GMM estimator simplifies to the Pearson correlation for bivariate cycle pairs whereas the multivariate synchronization indices stand for the restricted value of the bivariate Pearson correlation in higher dimensions, provided that the null hypothesis  $\text{SMS}(\rho_0)$  is not rejected, i.e.,  $W(\hat{\rho}_0) \leq CV_W$  (95%). The value for the test statistic and the critical value  $CV_W$  are also reported in the table.<sup>19</sup> The critical value is determined using a block bootstrap (see the appendix for further details on that procedure). Finally, the reported closed intervals  $[\rho_-, \rho_+]$  contain all values of  $\rho_0$  that lead to nonrejection of the null hypothesis of  $\text{SMS}(\rho_0)$ . In order to better grasp the results in Table 4.3, consider for example the index of multivariate synchronization  $\hat{\rho} = 0.42$  for the triplet  $[SING, THAI, KOR]$ . The bivariate correlations for this triplet are of the same order of magnitude indeed. Thus, it should not be surprising that the null hypothesis  $\rho[SING, THAI] = \rho[SING, KOR] = \rho[KOR, THAI]$  cannot be rejected over the interval  $[0.15, 0.69]$  which contains all three bivariate correlations. The nonrejection justifies the “restricted” 0.41 estimate and is obtained by minimizing the W-test over this interval.

If one compares the magnitude of synchronization indices in Table 4.3, the bivariate comovements seem to differ quite a lot at first sight, ranging from -0.04 (THAI, TAI) to 0.659 (all markets). However, the null hypothesis of a common stock market cycle ( $\text{SMS}(\rho_0)$ ) can nearly never be rejected which justifies the reported multivariate synchronization estimates in the lower panels of Table 4.3.

---

<sup>18</sup>For example, Korean and Thai stock markets already were in a bear phase prior to the Asian crisis for a considerable amount of time (three and one and a half years, respectively), see Edwards et al. (2003) and our own calculations.

<sup>19</sup>Because it only makes sense to test for multivariate synchronization when  $n > 2$ , the columns for the W-test and the critical values are left empty in the bivariate panel.

Those cases for which stock market cycles do not exhibit a common (homogeneous) correlation are denoted by “rej”. Notice also that the polar cases  $\rho = 0$  and 1 fall in the nonrejection intervals  $[\rho_-, \rho_+]$  for only seven out of 26 cases. This provides additional justification to allow for an “intermediate” estimation and testing procedure for “imperfect” multivariate synchronization.

Table 4.3: Full sample market synchronization: Estimates and homogeneity tests

country sets	$\hat{\rho}$	$[\hat{\rho}_-, \hat{\rho}_+]$	$W(\hat{\rho})$	$CV_W$
Panel A: n=2				
<i>SING; THAI</i>	0.43	[0.18, 0.68]	-	-
<i>SING; KOR</i>	0.47	[0.15, 0.79]	-	-
<i>SING; TAI</i>	0.34	[0.03, 0.66]	-	-
<i>SING; MAL</i>	0.33	[0.02, 0.65]	-	-
<i>THAI; KOR</i>	0.31	[0.03, 0.59]	-	-
<i>THAI; TAI</i>	-0.04	[-0.36, 0.27]	-	-
<i>THAI; MAL</i>	0.43	[0.16, 0.69]	-	-
<i>KOR; TAI</i>	0.40	[0.33, 1]	-	-
<i>KOR; MAL</i>	0.26	[0.23, 1]	-	-
<i>TAI; MAL</i>	0.22	[-0.10, 0.54]	-	-
Panel B: n=3				
<i>SING; THAI; KOR</i>	0.42	[0.15, 0.69]	2.81	14.14
<i>SING; THAI; TAI</i>	-	-	26.51	18.92
<i>SING; THAI; MAL</i>	0.42	[0.12, 0.72]	0.81	18.71
<i>SING; KOR; TAI</i>	0.41	[0.10, 0.73]	0.68	14.25
<i>SING; KOR; MAL</i>	0.38	[0.03, 0.74]	2.79	18.40
<i>SING; TAI; MAL</i>	0.31	[-0.07, 0.69]	1.39	18.78
<i>THAI; KOR; TAI</i>	-	-	24.58	15.44
<i>THAI; KOR; MAL</i>	0.37	[0.07, 0.66]	1.33	18.34
<i>THAI; TAI; MAL</i>	0.32	[0.14, 0.51]	12.99	19.87
<i>KOR; TAI; MAL</i>	0.33	[-0.01, 0.67]	2.31	18.18
Panel C: n=4				
<i>SING; THAI; KOR; TAI</i>	0.61	[0.52, 0.70]	31.96	36.79
<i>SING; THAI; KOR; MAL</i>	0.42	[0.09, 0.75]	6.29	35.33
<i>SING; THAI; TAI; MAL</i>	0.53	[0.39, 0.67]	30.58	38.44
<i>SING; KOR; TAI; MAL</i>	0.39	[0.00, 0.79]	5.35	32.23
<i>THAI; KOR; TAI; MAL</i>	0.50	[0.32, 0.69]	26.20	42.73
Panel D: n=5				
ALL	0.60	[0.36, 0.84]	36.78	81.90

**Note:** The estimated common synchronization index is denoted by  $\hat{\rho}$ . The closed interval  $[\rho_-, \rho_+]$  is the corresponding 95% confidence interval for  $\hat{\rho}$ .  $W(\hat{\rho})$  tests for the *SMS*( $\hat{\rho}$ ) hypothesis (test of “multivariate synchronization” or “homogeneity”).  $CV_W$  stands for the 95% critical value of the bootstrap version of the test. The bootstrap is performed on the binary series in blocks of length 25. In the bivariate case  $n = 2$ , tests for multivariate synchronization are meaningless because there is only one bivariate correlation.

We previously argued that cycle correlations - like the ones in Table 4.3 - are the more relevant risk diversification indicators, provided investors base their buying and selling decision on how stock cycles evolve (i.e., if investors look at cycle turning points to time their buying and selling decisions). Moreover, the conven-



tional correlations based on return pairs are found to be quite different for the majority of considered stock market pairs (not reported in the table). In other words, return correlations can provide misleading information about the potential for risk diversification when investors' time horizon (and thus their portfolio rebalancing) coincides with the stock market cycle.

We also want to know whether the cycle correlations in Table 4.3 are stable over the whole sample period. We earlier established (see Table 4.1 in the simulation section) that the bootstrapped versions of the restricted and unrestricted test perform equally well in eliminating the size distortion. We therefore opt for the simpler unrestricted test procedure (the bivariate correlations in the parameter vector in (4.6) are left unrestricted). Table 4.4 reports estimated break dates and corresponding values of the test statistic and the 95% critical value. Again, critical values  $CV_S$  are bootstrapped using the block bootstrap algorithm described in the appendix. As expected, the small sample critical values increase with the number of countries. This is consistent with the observation we made in the Monte Carlo section that size distortions of the asymptotic test increased with the number of countries. Moreover, the table shows that instability is generally present, both in bivariate and multivariate synchronization indices and that the majority of the breaks coincide with the Asian crisis era.<sup>20</sup>

Table 4.4 suggests that one should be very careful in interpreting the full sample correlations in Table 4.3 because they hide different subsample values. We therefore also calculated these corresponding subsample values in order to find out the direction of the change in synchronization (the two-sided stability test does not provide us with that information). The subsample synchronization indices  $\hat{\rho}$  for the subsamples defined in Table 4.4 are reported in the final Table 4.5. Table 4.5 also contains subsample testing outcomes for the null hypothesis of imperfect multivariate synchronization using (4.4). If this subsample application of the W-test does not lead to rejections, we are allowed to calculate multivariate synchronization indices for the subsamples. We do not report subsample results for those markets whose multivariate synchronization index does not exhibit breaks according to the preceding table (see "no breaks" rows).

The most striking table feature is that the synchronization index in the second subsample is nearly always much bigger than its pre-break counterpart.

---

<sup>20</sup>If e.g. financial liberalization, the Asian crisis, or changes in exchange rate regimes are the main triggers of the synchronization breaks in Asia, one would expect a comparable importance of breaks in Latin American synchronization and a much smaller number of breaks for developed stock markets. We therefore also calculated the stability test for these two "control" groups of stock markets. The Latin American panel consists of Argentina, Brazil, Mexico and Chile whereas the European panel contains the UK, Germany, France, Italy, the Netherlands, Belgium and Spain. We found significant breaks for nearly all Latin American synchronization combinations. In contrast, European synchronization was found to be remarkably stable.

Table 4.4: Testing for structural change in multivariate synchronization index

country sets	$\hat{b}(m/yy)$	$\sup(\Sigma_T(\hat{b}))$	$CV_S$
Panel A: n=2			
<i>SING; THAI</i>	no break	7.91	11.94
<i>SING; KOR</i>	4/92	37.53	13.28
<i>SING; TAI</i>	8/97	18.13	12.03
<i>SING; MAL</i>	no break	2.97	11.21
<i>THAI; KOR</i>	10/94	18.86	12.32
<i>THAI; TAI</i>	8/97	23.38	11.76
<i>THAI; MAL</i>	no break	6.75	12.76
<i>KOR; TAI</i>	8/97	26.97	13.16
<i>KOR; MAL</i>	3/97	18.47	11.87
<i>TAI; MAL</i>	8/97	22.68	11.76
Panel B: n=3			
<i>SING; THAI; KOR</i>	7/92	171.24	80.45
<i>SING; THAI; TAI</i>	8/97	141.32	82.28
<i>SING; THAI; MAL</i>	no break	28.00	84.09
<i>SING; KOR; TAI</i>	8/97	176.38	90.76
<i>SING; KOR; MAL</i>	7/92	196.61	87.16
<i>SING; TAI; MAL</i>	8/97	220.95	80.49
<i>THAI; KOR; TAI</i>	8/97	209.23	80.61
<i>THAI; KOR; MAL</i>	10/94	87.72	84.34
<i>THAI; TAI; MAL</i>	8/97	227.97	82.06
<i>KOR; TAI; MAL</i>	8/97	250.17	79.98
Panel C: n=4			
<i>SING; THAI; KOR; TAI</i>	-	-	-
<i>SING; THAI; KOR; MAL</i>	9/92	391.94	341.42
<i>SING; THAI; TAI; MAL</i>	10/97	508.59	326.04
<i>SING; KOR; TAI; MAL</i>	8/97	715.33	356.26
<i>THAI; KOR; TAI; MAL</i>	8/97	713.89	353.77
Panel D: n=5			
ALL	-	-	-

**Note:** The nominal size of the stability test is set to 5%. Table entry  $n$  refer to the number of countries.  $\sup(\Sigma_T(\hat{b}))$  is the statistic associated with the structural break test and  $CV_S$  is the corresponding bootstrapped critical value at 95%. The bootstrap is performed on the binary series in blocks of length 25. The date of break is indicated in the column  $\hat{b}(mm/yy)$ . Those cases in which the stability test could not be adequately performed due to the near-singular character of the variance-covariance matrix are denoted by (-).

Table 4.5: Subsample cyclical stock market synchronization in South East Asia: Estimates and homogeneity tests

country sets	subsample 1			subsample 2		
	$\hat{\rho}_1$	$[\hat{\rho}_-, \hat{\rho}_+]$	$W(\hat{\rho}_1)$	$CV_W$	$\hat{\rho}_2$	$W(\hat{\rho}_2)$
Panel A: n=2						
<i>SING; THAI</i>				no break		
<i>SING; KOR</i>	-0.26	$[-0.55, 0.03]$			0.76	$[0.61, 0.92]$
<i>SING; TAI</i>	-0.01	$[-0.46, 0.44]$			0.74	$[0.48, 1]$
<i>SING; MAL</i>				no break		
<i>THAI; KOR</i>	-0.21	$[-0.48, 0.08]$			0.54	$[0.26, 0.82]$
<i>THAI; TAI</i>	-0.46	$[-0.53, -0.39]$			0.45	$[0.06, 0.85]$
<i>THAI; MAL</i>				no break		
<i>KOR; TAI</i>	0.04	$[-0.39, 0.46]$			0.87	$[0.73, 1]$
<i>KOR; MAL</i>	-0.16	$[-0.41, 0.08]$			0.62	$[0.23, 1]$
<i>TAI; MAL</i>	-0.21	$[-0.49, 0.06]$			0.65	$[0.24, 1]$
Panel B: n=3						
<i>SING; THAI; KOR</i>	-0.14	$[-0.47, 0.18]$	2.58	41.49	0.73	$[0.51, 0.95]$
<i>SING; THAI; TAI</i>	rej.		116.84	27.17	0.70	$[0.32, 1]$
<i>SING; THAI; MAL</i>				no break		
<i>SING; KOR; TAI</i>	0.08	$[-0.30, 0.46]$	0.74	31.38	0.89	$[0.66, 1]$
<i>SING; KOR; MAL</i>	-0.01	$[-0.14, 0.13]$	10.56	29.50	0.84	$[0.69, 1]$
<i>SING; TAI; MAL</i>	-0.06	$[-0.30, 0.16]$	3.91	24.95	0.96	$[0.80, 1]$
<i>THAI; KOR; TAI</i>	rej.		58.11	29.82	0.85	$[0.65, 1]$
<i>THAI; KOR; MAL</i>	0.03	$[-0.05, 0.12]$	38.09	40.49	0.61	$[0.38, 0.83]$
<i>THAI; TAI; MAL</i>	rej.		160.27	25.73	0.77	$[0.49, 1]$
<i>KOR; TAI; MAL</i>	-0.09	$[-0.38, 0.19]$	2.40	25.87	0.85	$[0.58, 1]$
Panel C: n=4						
<i>SING; THAI; KOR; TAI</i>	-		-	-	-	-
<i>SING; THAI; KOR; MAL</i>	0.11	$[-0.07, 0.29]$	85.61	131.92	0.86	$[0.75, 0.97]$
<i>SING; THAI; TAI; MAL</i>	rej.		235.99	78.43	0.92	$[0.66, 1]$
<i>SING; KOR; TAI; MAL</i>	0.02	$[-0.25, 0.3]$	12.56	77.35	0.94	$[0.79, 1]$
<i>THAI; KOR; TAI; MAL</i>	rej.		336.52	76.39	0.85	$[0.5, 1]$

**Note:** Variables indexed by 1 (resp. 2) refer to the pre- (resp. post) break period.  $\hat{\rho}$  represents estimates of the synchronization index and  $[\rho_-, \rho_+]$  is the corresponding confidence interval at 95%.  $W(\hat{\rho})$  is the test statistic of the  $SM S(\hat{\rho})$  hypothesis and  $CV_W$  is the 95% critical value of the bootstrap version of the test. The bootstrap is performed on the binary series in blocks of length 25. Those cases in which the stability test could not be adequately performed due to the near-singular character of the variance-covariance matrix are denoted by (-). This is for the sample with all countries  $n = 5$  the case and the reason why we refrain from showing those results in a separate panel. In the bivariate case  $n = 2$ , tests for multivariate synchronization are meaningless because there is only one bivariate correlation.

Moreover, the pre-break nonrejection intervals are lower and do not overlap with the post-break nonrejection intervals. Both observations are consistent with the breakpoint outcomes of the previous table. The confidence intervals actually suggest that pre-break synchronization was often insignificantly different from zero whereas post-break synchronization was close to perfect ( $\rho = 1$ ) for more than half of the cases. As expected, the full sample synchronization index lies somewhere in between the subsample values and can be considered as a rough average of the true subsample values. Turning to the outcomes of the subsample homogeneity test, we find only a limited number of rejections for the first subsample; but deviations from homogeneity all disappear in the second subsample. Thus, the multivariate synchronization index increases over time and heterogeneity - present in the first subsample on a limited scale - almost completely disappears after the breaks.

The observed increases in cyclical correlations suggest that the room for risk diversification drastically diminished after the Asian crisis (at least for portfolio managers that solely invest in the considered Asian markets and who buy and sell according to the turning points of the stock market cycle). If the increased synchronization persists in the long run (i.e. a permanent change), then the investors' (long run) "strategic" asset allocation will be affected, whereas if the rise is transitory, it will only affect their "tactical" asset allocation (portfolio composition in the short run). Insofar as the rise in stock market synchronization is permanent, regulatory authorities in the different countries probably have to adjust financial regulation in order to preserve banking system stability in the Asian financial system. More specifically, a high cross country stock market synchronization can lead to boom-bust credit cycles spilling over from one country to the other.

On the other hand, if the increased synchronization is a purely contagious and short run phenomenon, policy makers cannot do much more than (i) to avoid these spillovers by preventing the development of boom-bust cycles in their domestic economies and (ii) to mitigate the financial and real effects of the contagion if bulls and bears spill over to other markets.

Whether the increase in Asian stock market synchronization has a permanent or transitory character is open to debate. As a matter of fact, our current econometric framework is not able to disentangle whether the increase in stock market synchronization is permanent or transitory. Indeed, notice that the sheer magnitude of the subsample changes  $\hat{\rho}_2 - \hat{\rho}_1$  is comparable to the largest synchronization changes assumed in the Monte Carlo power study. The magnitude of the jump and the inability to detect gradual breaks (see simulation section) suggest that the breaks and corresponding subsample results in Tables 4.4 and 4.5 provide evidence against financial integration. However, the observed sudden rises in synchronization are not necessarily interpretable as evidence pro "financial contagion". The

latter phenomenon would require, *inter alia*, that the jump in the correlations is only temporary.

## 4.5 Conclusions

In this chapter we proposed a Generalized Method of Moments (GMM) framework to measure the degree of synchronization between stock market “bulls” and “bears”. We argued that an assessment of cycle duration and cycle comovement is potentially relevant for investors that base their investment decisions on the turning points of the stock market cycle. Moreover, policy makers and regulators might be interested to know the magnitude of stock market synchronization and whether it changed over time because of the potentially destabilizing effects for stock market bulls and bears on the real economy.

Prior to calculating a measure of cyclical synchronization, we classified stock prices into “bull” and “bear” periods using the Bry and Boschan (1971) dating algorithm. We subsequently extended the Harding and Pagan (2006b) framework in several directions.

First, we allowed for a value of the common synchronization index to be between -1 and 1 whereas Harding and Pagan (2006b) only tested against the benchmark cases of complete perfect synchronization or nonsynchronization (so, they did not really estimate a common synchronization index but restricted it to either 0, -1 or +1 prior to performing the test for a common cycle). However, in practice, business cycles as well as financial cycles are neither perfectly synchronized nor completely independent which implies that our testing framework is closer to reality than Harding and Pagan’s. Moreover, our approach also produces an estimate of the multivariate synchronization index  $\rho_0$  ( $-1 < \rho_0 < 1$ ).

Second, we proposed an endogenous stability testing procedure for detecting structural change in the cyclical stock market synchronization index  $\rho_0$ . Before putting the test to work in an empirical application, we performed a Monte Carlo experiment to evaluate the (small sample) size and power properties of the novel testing procedure. We found that the stability test suffers from massive overrejection when one uses asymptotic critical values and when the number of stock markets considered grows large. However, a bootstrap of the small sample distribution can remedy this problem fairly easily. In the power study, we made use of the small sample critical values to obtain size-adjusted power values. We found that the stability test is able to detect breaks reasonably well provided that the changes in synchronization occur suddenly and are relatively large in magnitude. Indeed, the power deteriorates surprisingly quickly upon lowering the adjustment speed or decreasing the change in the synchronization index. In other words, the

stability test seems unable to pick up gradual structural breaks in synchronization which means that a financial integration interpretation for breaks is likely to be wrong in our framework. It is then tempting to interpret the breaks as evidence for financial contagion. However, one should be cautious with that break interpretation because the stability test is unable to distinguish permanent shocks from transitory shocks and contagion is by definition a transitory phenomenon.

Changes in the synchronization of stock market cycles are probably more likely to occur across emerging markets than across developed markets due to financial liberalization and recurrent financial turmoil. We therefore selected a set of Asian stock markets for our empirical application. Upon applying the stability test, we detected an increase in synchronization, mainly after the Asian crisis, that is both economically and statistically significant. Upon applying the test for multivariate synchronization on the subsamples defined by the breaks, the pairwise (bivariate) synchronization indices seem to converge even more towards each other after the break. Moreover, we were unable to find breaks for a control group of developed countries which seems to confirm that forces like financial liberalization, institutional reform and market turbulence like the Asian crisis - that have been less prominent in developed markets - might be responsible for the increased synchronization.

The observed rise in Asian stock market synchronization implies that there is less space for diversifying risk after the Asian crisis (at least for investors that solely invest in the considered Asian markets and whose portfolio rebalancing is dictated by the turning points of the stock market cycle). If the rise in stock market synchronization has a lasting character, regulatory bodies probably need to change their supervisory framework in order to preserve banking system stability in the Asian financial system. On the other hand, if the stronger comovements between bulls and bears is a purely transitory (and possibly contagious) phenomenon, policy makers cannot do more than (i) preventing these spillovers by reducing the potential for the build up of boom-bust cycles in their domestic economies and (ii) reducing the financial and real effects of the transitory shock if bulls and bears spill over to other markets.

The observed post-break increases in synchronization possibly contain a permanent as well as a transitory component. First, one could imagine that the Asian crisis and its direct aftermath had a contagious character as many authors have claimed since then. Subsequently, policymakers and regulatory bodies implemented a myriad of measures but the recipes for both dampening the effects of the Asian crisis and reducing the potential of a future crises to strike and spread across borders were pretty similar in all affected countries. This “convergence” in post-crisis policy measures might itself have had a long run impact on the

synchronization correlation. Disentangling the observed rise on synchronization correlations into a permanent and a transitory effect makes part of our future research agenda.

## 4.6 Appendix: Block bootstrap algorithm

Koenker and Machado (1999) argued that size distortions due to a large number of GMM moment conditions can be partly remedied by increasing the sample size  $T$ . However, this does not seem to be a realistic option in our case; our own simulations (not included in the tables) have shown that one needs time series of several thousands of observations in order to mitigate the size distortion. We therefore opt to bootstrap the small sample distribution of the test statistics (4.3) and (4.6). The asymptotic pivotality of both tests and the fact that they are continuous functions of sample moments are sufficient to ensure proper convergence of the bootstrapped critical values to the true asymptotic distribution, see e.g. Horowitz (2001).

Although suitable regularity conditions guarantee proper asymptotic convergence of the bootstrap, we also want to know whether the bootstrap works in small samples, i.e., whether the bootstrapped critical values are sufficiently size-corrected. We limit ourselves to an exposition of the bootstrap algorithm for the critical values of the multivariate synchronization test in (4.3), the procedure for the stability test being completely analogous. For sake of convenience, we limit ourselves to study the performance of the bootstrap under the null hypothesis of Strong Multivariate Nonsynchronization (SMNS), i.e.,  $H_0 : \rho_0 = 0$ . (The same null was assumed for the stability test bootstrap). We set up a Monte Carlo cum bootstrap experiment that consists of the following steps:

1. Step MC1: Assume a particular DGP and a filter  $\varphi(\cdot)$ . Draw a  $T \times n$  matrix by using either  $DGP_1$  (IID returns) or  $DGP_2$  (Bollerslev's (1990) Constant Conditional Correlation model). This renders the  $T \times n$  matrix  $X$ .
2. Step MC2: Use the Bry and Boschan (1971) dating algorithm to build a  $T \times n$  binary variable matrix  $S$  with characteristic element  $\varphi(X_{i,t}^{MC}) = S_{i,t}^{MC}$ .
3. Step MC3: Compute the simulated W-statistic for  $\rho_0 = 0$ , i.e.,  $W^{MC}(\rho_0 = 0)$  using  $S$  as input.
4. Step BO1: Generate a  $T \times n$  matrix of bootstrap replications  $S^B$  with block length 25 in the time series dimension.
5. Step BO2: Compute the bootstrapped W-statistic for  $\rho_0 = 0$ , i.e.,  $W^B(\rho_0 = 0)$  using  $S^B$  as input.
6. Repeat the bootstrap steps 4-5 a "sufficient" number of times ( $M$ ).  $M$  is endogenously determined using the three step-method of Andrews and



Buchinsky (2000).<sup>21</sup> A critical value is then obtained as the  $\alpha$  percent quantile, say  $W_{crit}^B$ , from the empirical distribution of the bootstrap test statistic. The nominal size is set to  $\alpha = 5\%$ . The null hypothesis of  $SMS(\rho_0 = 0)$  is rejected if  $W^{MC}(\rho_0 = 0) > W_{crit}^B$ .

7. Repeat steps *MC1* to *MC3*, and *BO1* to *BO3* a large number of times to obtain the rejection frequencies of the test  $W(\rho_0 = 0)$ .

We opted for bootstrapping in blocks (BO1) because we know that the binary time series  $S_{i,t}$  ( $i = 1, \dots, n$ ) that constitute the columns of  $S$  can be described by higher order stationary and ergodic Markov chains, see Harding and Pagan (2006a, p.62). The determination of the optimal block length is an ongoing point of discussion in the bootstrap literature. Hall et al. (1995) suggest to set  $w = c.T^{1/4}$ , with  $c$  a fixed parameter depending on the DGP. We found by trial and error that the size distortion is best tackled with block lengths of size 25.

---

<sup>21</sup>The endogenous number of bootstrap replications  $M$  was found to vary between 200 and 2,000.

---

## COMOVEMENTS OF DIFFERENT ASSET CLASSES DURING MARKET STRESS

Here we assess the linkages between the most important U.S. financial asset classes (stocks, bonds, T-bills and gold) during periods of financial turmoil.<sup>1</sup> Our results have potentially important implications for strategic asset allocation and pension fund management. We use multivariate extreme value theory to estimate the exposure of one asset class to extreme movements in the other asset classes. By applying structural break tests to those measures we study to what extent linkages in extreme asset returns and volatilities are changing over time. Univariate results and bivariate comovement results exhibit significant breaks in the 1970s and 1980s corresponding to the turbulent times of e.g. the oil shocks, Volcker's presidency of the Fed or the stock market crash of 1987.

### 5.1 Introduction

Joint crashes in different asset markets can have severely destabilizing effects on countries and the international financial architecture. Strong financial market linkages during crisis periods can severely increase the risk of bank failures through a joint deterioration of their assets possibly leading to domino effects in countries' financial systems even when their banks' assets are well diversified. In general, extreme comovements of financial markets on a national and international level crucially determine the systemic risk of these markets. The amount and size of jointly affected markets together with potential difficulties and bottlenecks in the financial system and payment process determines the severeness of any real effects that may follow. There have been periods with financial and political instability like the oil crises in 70s and 80s, the Asian Flu and Russian Cold (1997 and 1998, respectively) or more recently the subprime mortgage crises in 2007, where such

---

<sup>1</sup>This chapter is based on the paper Piplack and Straetmans (2008).

effects have been witnessed. Thus, the study of extreme (co)movements in asset markets is not only important to investors but also to policy makers and financial regulators that care about overall economic and financial stability.

Possibly the first systematic study of cross-country financial crisis spillovers is Morgenstern (1956, Chapter X). He explicitly refers to “statistical extremes” of the 23 stock markets and their effects on foreign stock markets. More recently, the econometric literature utilizes correlation analysis based, for example, on ARCH and GARCH-type models. Such contributions usually examine if stock market comovements differ between crisis and non-crisis episodes and typically also try to determine the direction of possible spill-over effects. Contributions like King and Wadhvani (1990), Hamao et al. (1990), Mallaris and Urrutia (1992), Lin et al. (1994) and Engle and Susmel (1993) belong to this strand of the literature. Papers focussing on foreign exchange markets and currency crises include Eichengreen et al. (1996), Sachs et al. (1996), Kaminski and Reinhart (2000). However, little work has been done on linkages across asset classes. Hartmann et al. (2004) constitutes a notable exception.

This chapter extends the literature by increasing the amount of asset classes considered. This allows us to study and compare phenomena like “flight to quality”, and “flight to liquidity”. We define flight to quality as the simultaneous event of a stock market crash and a boom in either government bond or gold markets; whereas flight to liquidity stands for a stock market crash coinciding with a boom in the market for T-bills. Compared to the scant existing literature on cross-asset linkages, we use more assets and longer time series. This allows us to implement extreme value techniques and to apply tests for structural change on our linkage measures.

The used methodology combines extreme value theory (EVT) with a structural stability test developed by Quintos et al. (2001). Contributions using similar approaches include Hartmann et al. (2004; 2005), and Straetmans et al. (2006). Bivariate extreme value theory captures the dependence structure in the tails of multivariate distributions by means of the so-called tail dependence parameter. This parameter is able to capture both linear and nonlinear dependence in the tails whereas traditional correlation analysis only measures linear dependence and is predisposed toward the multivariate normal distribution. Another advantage constitutes the nonparametric character of the used methodology, i.e., we leave the joint asset return process unspecified and thereby limit the scope for misspecification (model risk).

Anticipating on our results, we find relatively small tail indexes for gold and T-bills as compared to stocks and bonds. Bivariate results indicate that the likelihood of co-crashes dominates flight to quality and flight to liquidity phenomena.

As concerns structural change, both univariate and bivariate tails are found to be nonstable over time for certain asset pairs. The breaks suggest a mean reverting pattern in the amount of tail thickness and tail dependence: initially the probability mass has risen (oil shocks) to decline later on towards the end of the 80s. Tail asymmetries as well as cross sectional differences in tail estimates are found to be statistically insignificant from zero.

The chapter is organized as follows. Section 5.2 introduces the theoretical basis for the extreme value analysis. Section 5.3 explains the tail-dependence measure for extreme financial market comovements in more detail. In Section 5.4 we introduce the stability test that we perform in order to check for structural breaks in the univariate and multivariate series. Section 5.5 presents the results obtained by applying those techniques to the data. Section 5.6 concludes.

## 5.2 Asset linkages: Theory

We measure the dependencies between returns of different asset classes for extreme price movements, i.e. in the bivariate tail of the asset pairs. They are constructed either as conditional tail probabilities or conditionally expected extreme co-events. We will argue that the two indicators are perfectly correlated and are two alternative ways to presenting the same empirical outcomes. The techniques used are not new and have partly been used in, for example, Poon et al. (2004) and Hartmann et al. (2004; 2005).

In this section and Section 5.3 we assume constancy/stationarity of the tail behavior of assets over time. We also focus on the unconditional marginal return distributions and do not condition any statistic on time. Therefore, we refrain from using time subscripts even though the reader should bear in mind that the assumed asset return series evolve over time. In Section 5.4 we introduce time subscripts  $t$  because we relax the assumption of constancy of the tail behavior and allow (test) for structural breaks.

### 5.2.1 Conditional tail probabilities

Consider a pair of different asset types, i.e., stocks and bonds. Denote the return of stocks and bonds by the random variables  $X_i$  ( $i = 1, 2$ ), respectively. Each series  $X_i$  is assumed to have  $n$  observations. For sake of convenience and when necessary, we take the negative of returns, so that we can define all used formulae in terms of upper tail returns. Crisis levels or extreme percentiles  $Q_i$  ( $i = 1, 2$ ) are chosen such that the tail probabilities are equal across assets, i.e.,  $P\{X_1 > Q_1\} = P\{X_2 > Q_2\} = p$ .

With common marginal exceedance probabilities, crisis levels  $Q_i$  (Value-at-Risk/VaR) will generally not be the same across assets, because the marginal distribution functions  $P\{X_i > Q_i\} = 1 - F_i(Q_i)$  are nonidentical. Crisis levels can be interpreted as ‘barriers’ that will on average only be broken once in  $1/p$  time periods, i.e.,  $p^{-1}$  days in case of daily data frequency. Suppose now that we want to measure the dependence between two assets beyond the crisis levels  $(Q_1, Q_2)$ . A natural measure is the conditional tail probability

$$\begin{aligned}\beta_\tau &: = P\{X_1 > Q_1(p) | X_2 > Q_2(p)\} \\ &= \frac{P\{X_1 > Q_1(p), X_2 > Q_2(p)\}}{P\{X_2 > Q_2(p)\}} \\ &= \frac{P\{X_1 > Q_1(p), X_2 > Q_2(p)\}}{p},\end{aligned}\tag{5.1}$$

which measures the likelihood that an asset’s value (in this case  $X_1$ ) falls sharply, if there is an extreme negative shock to a second asset. In case of independence the conditional tail probability reduces to  $p^2/p = p$ , which constitutes a lower bound that helps to judge the strength of assets’ tail dependence.

### 5.2.2 Conditionally expected extreme events

Alternatively, suppose we would like to find the expected number of assets’ extremes (booms or busts) given that one observes a boom or bust in at least one asset class. Using the same notation as before, we represent random asset returns by  $X_1$  and  $X_2$ .  $Q_1$  and  $Q_2$  are the corresponding percentiles (or ‘thresholds’) above which we speak of a market boom or crash (in case of a loss) and that will only be exceeded with probability  $p$ . Let  $\kappa$  stand for the number of assets with extreme returns, i.e.  $\kappa$  equals one or two. Our extreme linkage indicator is the conditional expectation  $E[\kappa | \kappa \geq 1]$ . From elementary probability theory (starting from the standard definition of conditional probability) we can state that

$$\begin{aligned}E[\kappa | \kappa \geq 1] &:= \frac{E[\kappa]}{P\{\kappa \geq 1\}} \\ &= \frac{P\{X_1 > Q_1, X_2 \leq Q_2\} + P\{X_1 \leq Q_1, X_2 > Q_2\}}{P\{X_1 \geq Q_1 \text{ or } X_2 \geq Q_2\}} \\ &\quad + \frac{2P\{X_1 > Q_1, X_2 > Q_2\}}{P\{X_1 \geq Q_1 \text{ or } X_2 \geq Q_2\}} \\ &= \frac{2p}{P\{X_1 \geq Q_1 \text{ or } X_2 \geq Q_2\}}\end{aligned}\tag{5.2}$$

with  $P\{X_1 \geq Q_1 \text{ or } X_2 \geq Q_2\} = 1 - P\{X_1 \leq Q_1, X_2 \leq Q_2\}$ . Notice that the conditional expectation reduces to  $2/(2-p)$  under the benchmark of independence. It is also easily observed that  $E[\kappa|\kappa \geq 1] = P\{\kappa = 2|\kappa \geq 1\} + 1$ , so that an alternative interpretation of our extreme linkage indicator is in terms of (1 plus) the conditional probability that both assets simultaneously boom or bust given that at least one asset exhibits extreme behavior. For higher dimensions than two  $E[\kappa|\kappa \geq 1]$  is still equal to the ratio of the sum of the marginal excess probabilities divided by the joint failure probability. The relation between both extreme linkage measures (5.1) and (5.2) easily follows from the following chain of equalities:

$$\begin{aligned} E[\kappa|\kappa \geq 1] &= \frac{2p}{P\{X_1 \geq Q_1 \text{ or } X_2 \geq Q_2\}} \\ &= \frac{2p}{2p - P\{X_1 \geq Q_1, X_2 \geq Q_2\}} \\ &= \frac{2}{2 - \beta_\tau}. \end{aligned}$$

Clearly,  $1 \leq E \leq 2$  corresponds with  $0 \leq \beta_\tau \leq 1$ .

### 5.3 Estimation of the linkage indicators

The estimation of (5.1) and (5.2) reduces to the estimation of the joint probability  $P\{X_1 \geq Q_1, X_2 \geq Q_2\}$ . Within the framework of a parametric probability law, the calculation of the proposed multivariate probability measures is straightforward, because one can estimate the distributional parameters by, e.g., maximum likelihood techniques. However, if one makes the wrong distributional assumptions, the linkage estimates may be severely biased due to misspecification. As there is no clear evidence that all asset returns follow the same distribution – even less so for the crisis situations we are interested in here – we want to avoid very specific assumptions for assets' returns. Therefore, we implement the semi-parametric EVT approach proposed by Ledford and Tawn (1996); see also Draisma et al. (2001), and Poon et al. (2004) for recent applications). Loosely speaking, their approach consists of generalizing some 'best practice' in univariate extreme value analysis.

Before proceeding with the modeling of the extreme dependence structure, however, it is worthwhile to eliminate any possible influence of marginal aspects on the joint tail probabilities by transforming the original variables to a common marginal distribution. After such a transformation, differences in joint tail probabilities can be solely attributed to differences in the tail dependence structure of the extremes. Thus our dependence measures, unlike e.g. correlation, are no longer influenced by the differences in marginal distribution shapes. To this aim

we transform asset returns  $(X_1, X_2)$  to unit Pareto marginals:

$$\tilde{X}_i = \frac{1}{1 - F_i(X_i)}, \quad i = 1, 2, \quad (5.3)$$

with  $F_i(\cdot)$  representing the marginal cumulative distribution function (cdf) for  $X_i$ .<sup>2</sup> This variable transform leaves the joint tail probability in the numerator of (5.1) invariant because

$$P\{X_1 > Q_1(p), X_2 > Q_2(p)\} = P\{\tilde{X}_1 > s, \tilde{X}_2 > s\},$$

with  $s = 1/p$ .<sup>3</sup> The estimation problem can now be simplified toward estimating a univariate exceedance probability for the cross-sectional minimum of the two return series, i.e., it is always true that:

$$P\{\tilde{X}_1 > s, \tilde{X}_2 > s\} = P\{\min(\tilde{X}_1, \tilde{X}_2) > s\} = P\{Z_{\min} > s\} \quad (5.4)$$

The marginal tail probability at the right-hand side can now be easily calculated by making an additional assumption on the univariate tail behavior of  $Z_{\min}$ . Ledford and Tawn (1996) argue that the bivariate dependence structure is a regularly varying function under fairly general conditions. Draisma et al. (2001) give sufficient conditions and further motivation. Therefore, we assume that the auxiliary variable  $Z_{\min}$  has a regularly varying tail. An intuitive justification of the regular variation assumption for the bivariate tail lies in the generally observed regular variation (heavy tails or non-normality) of the original return series  $X_1$  and  $X_2$ . Thus, it is reasonable to assume that the transformed series  $\tilde{X}_i$  and hence the series of the cross-sectional minimum in (5.4) should also inherit this property. Upon assuming that  $Z_{\min}$  exhibits a fat tail, the regular variation assumption means that the marginal excess probability for the tail of the auxiliary variable in (5.4) has a Pareto tail decline:

$$P\{Z_{\min} > s\} \approx L(s) s^{-\alpha}, \quad \alpha \geq 1 \quad (5.5)$$

---

<sup>2</sup>Since  $F_{1,2}$  are unknown, we replace them with their empirical counterparts. For each  $X_i$  this leads (with a small modification to prevent division by 0) to:

$$\tilde{X}_i = \frac{1}{1 - R_{X_i}/(n+1)}, \quad i = 1, 2,$$

where  $R_{X_i} = \text{rank}(X_{ij}, j = 1, \dots, n)$ .

<sup>3</sup>The joint probability stays invariant under any monotonically increasing transformation of the marginals.

with  $s$  large ( $p$  small) and where  $L(s)$  is a slowly varying function.<sup>4</sup> Distributions with a Pareto-type tail decline have bounded moments only up to  $\alpha$ , where  $\alpha$  is the ‘tail index’ of  $Z_{\min}$ . In contrast, distributions with exponentially decaying tails (e.g. the normal df) or with finite endpoints have all moments bounded. So, the larger  $\alpha$  the thinner is the tail of a distribution.<sup>5</sup> We can now distinguish two cases in which the  $\tilde{X}_i$  ( $i = 1, 2$ ) are either *tail dependent* or *independent*. In the former case,  $\alpha = 1$  and

$$\lim_{s \rightarrow \infty} P \left\{ \tilde{X}_1 > s \mid \tilde{X}_2 > s \right\} > 0.$$

Stated otherwise, the tail probability defined on the pair of random variables  $(X_1, X_2)$  does not vanish in the bivariate tail. Examples of asymptotically dependent random variables include the multivariate student-t distribution and the multivariate logistic distribution, see e.g. Longin and Solnik (2001), Poon et al. (2004). For *asymptotic independence* of the random variables  $\alpha > 1$ , and we have that

$$\lim_{s \rightarrow \infty} P \left\{ \tilde{X}_1 > s \mid \tilde{X}_2 > s \right\} = 0.$$

Examples of this class of distributions include the bivariate standard normal distribution or the bivariate Morgenstern distribution. For the bivariate normal with nonzero correlation coefficient  $\rho$ , the auxiliary variable’s tail descent in (5.4) will be governed by  $\alpha = 2/(1 + \rho)$  whereas the bivariate Morgenstern corresponds with  $\alpha = 2$ . Notice that we only reach  $\alpha = 2$  for the bivariate standard normal when  $\rho = 0$ . In general, whenever the  $\tilde{X}_i$  ( $i = 1, 2$ ) are fully independent,  $\alpha = 2$  and  $P\{Z_{\min} > s\} = p^2$ . But the reverse is not true, i.e., there are joint distributions with nonzero pairwise correlation that nevertheless have  $\alpha = 2$ . The above-mentioned Morgenstern model provides an example. When the normal random variables are independent ( $\rho = 0$ ), the joint excess probability is also governed by  $\alpha = 2$ .

The steps (5.3), (5.4) and (5.5) show that the estimation of joint probabilities like in (5.4) can be reduced to a univariate estimation problem. Univariate excess probabilities can be estimated by using the semi-parametric probability estimator from De Haan et al. (1994):

$$\hat{p}_s = \frac{m}{n} (Z_{n-m,n})^\alpha s^{-\alpha}, \quad (5.6)$$

where the ‘tail cut-off point’  $Z_{n-m,n}$  is the  $(n - m)$ -th ascending order statistic (or loosely speaking the  $m$ -th smallest return with  $m$  being the amount of returns

<sup>4</sup>i.e.,  $\lim_{s \rightarrow \infty} L(ts)/L(s) = 1$  for all fixed  $t > 0$ .

<sup>5</sup>Such an interpretation holds for the univariate and for the multivariate case.



belonging to the tail of the distribution) of the auxiliary variable  $Z_{\min}$ .<sup>6</sup> Below we explain how we chose  $m$ .

The probability estimator (5.6) still needs a tail index estimate  $\alpha$  as an input. We estimate the tail index of the  $Z_{\min}$  series by means of the popular Hill (1975) statistic:

$$\hat{\alpha} = \left( \frac{1}{m} \sum_{j=0}^{m-1} \ln \left( \frac{Z_{n-j,n}}{Z_{n-m,n}} \right) \right)^{-1}, \quad (5.7)$$

where  $m$  has the same value and interpretation as in (5.6). Further details on the Hill estimator and related procedures to estimate the tail index are provided in Jansen and de Vries (1991) or the monograph by Embrechts et al. (1997).<sup>7</sup>

The above discussion demonstrates that the pair of estimators in (5.6)-(5.7) both characterizes univariate and multivariate tail behavior. This is because the estimation of a joint exceedance probability can be reduced to estimating a univariate exceedance probability. In the latter case, the tail index  $\alpha$  not only signals the tail thickness of the auxiliary variable  $Z_{\min}$  but it also reflects the strength of the dependence in the tails of the original return pair  $(X_1, X_2)$  in the tail area  $[Q_1, \infty) \times [Q_2, \infty)$ . The smaller the value of  $\alpha$  the higher the probability mass in the tail of  $Z_{\min}$  and thus also the higher the value of the joint probability in (5.1). One therefore often calls the inverse parameter  $\eta = 1/\alpha$  the tail dependence coefficient. An estimator of the bivariate tail probability measure in (5.1) now easily follows by combining (5.6) and (5.7):

$$\begin{aligned} \hat{\beta}_\tau &= \frac{\hat{p}_s}{p} \\ &= \frac{m}{n} (Z_{n-m,n})^{1/\hat{\eta}} s^{1-1/\hat{\eta}} \end{aligned} \quad (5.8)$$

for large but finite  $s = 1/p$ . When the original pair of returns exhibit asymptotic independence ( $\eta < 1$ ), the tail probability is a declining function of the threshold  $s$  and converges to zero if  $s \rightarrow \infty$ . On the other hand, in the polar case of asymptotic or tail dependence ( $\eta = 1$ ), the tail probability will always be above zero (regardless of the value of the conditioning percentile). However, in this chapter we will not focus on the asymptotic dependence vs. independence debate and will leave the tail dependence coefficient unrestricted. Moreover, Poon et al. (2004) already noticed that wrongly imposing asymptotic dependence ( $\eta = 1$ ), if the returns are

---

<sup>6</sup>Such a procedure can also be used for more than two return series as is done in, for example, Hartmann et al. (2005).

<sup>7</sup>Hill (1975) derived asymptotic consistency and normality of the Hill estimator under an i.i.d. assumption. Hsing (1991) and Resnick and Stărică (1998) derive similar results for the case of dependent data. For technical details we refer to the respective papers.

actually independent in the limit, might lead to severe overestimation of extreme linkage measures like (5.1). Thus, our approach is more flexible and we avoid the risk of overestimation.

Notice that the Hill statistic (5.7) still requires the choice of a nuisance parameter  $m$ , i.e., where do we let the tail start? Goldie and Smith (1987) suggest to select  $m$  such as to minimize the asymptotic mean-squared error (AMSE) of the Hill statistic in (5.7). Such a minimum should exist because of the bias-variance trade-off that is characteristic for the Hill estimator. This idea of balancing the bias and variance has become the cornerstone for most empirical techniques to determine  $m$ . We opted for the Beirlant et al. (1999) algorithm who proposed to use an exponential regression model (ERM) on the basis of scaled log-spacings between subsequent extreme order statistics from a Pareto-type distribution. Running Least Squares regressions on this exponential regression model allows one to estimate the AMSE for different  $m$ -values and to choose the optimal  $m$  that minimizes the AMSE. For more details on the algorithm we refer to the cited reference.

## 5.4 Hypothesis testing

In this section we introduce tests that can be used to assess various hypotheses regarding the temporal stability and cross sectional equality of the considered asset linkage indicators. The first one allows to test for the structural stability of the two indicators whereas the second test compares linkage indicators both across asset pairs and across time.

### 5.4.1 Time variation

The theory up to now assumed stationarity of tail behavior over time. From e.g. a strategic asset allocation perspective, however, it is important to know whether these interdependencies stay constant over time. As the discussion of the Ledford and Tawn (1996) approach toward estimating (5.1) has shown, the structural (in)stability of the indicators will critically depend on whether the tail dependence parameter  $\eta$  is constant or not. We therefore study possible temporal shifts in  $\eta$  with a recently developed structural stability test for the Hill statistic (5.7).

Quintos et al. (2001) present a number of tests for identifying a single unknown break in the estimated tail index  $\hat{\alpha}$ . As our estimation approach allows to map the multivariate dependence problem into a univariate estimation problem, we can choose from them the best test procedures for our tail dependence parameter  $\eta$ . Balancing the prevention of type I and type II errors we opt for their recursive

test.

Let  $t$  denote the endpoint of a sub-sample of size  $w_t < n$ . The recursive estimator for the tail dependence parameter  $\eta$  is calculated from (5.7) for sub-samples  $[1; t] \subset [1; n]$ <sup>8</sup>:

$$\hat{\eta}_t = \frac{1}{m_t} \sum_{j=0}^{m_t-1} \ln \left( \frac{Z_{t-j,t}}{Z_{t-m_t,t}} \right), \quad (5.9)$$

with  $m_t = \kappa t^{2/3}$ .<sup>9</sup>

The value of the recursive test statistic equals the supremum of the following series:

$$Y_n^2(t) = \left( \frac{tm_t}{n} \right) \left( \frac{\hat{\eta}_n}{\hat{\eta}_t} - 1 \right)^2. \quad (5.10)$$

Expression (5.10) compares the recursive value of the estimated tail parameter (5.7) with its full sample counterpart  $\hat{\eta}_n$ . The null hypothesis of interest is that the tail dependence parameter does not exhibit any temporal changes. More specifically, let  $\eta_t$  be the dependence in the left tail of  $Z$ .<sup>10</sup> The null hypothesis of constancy then takes the form

$$H_0 : \eta_{[nr]} = \eta, \quad \forall r \in R_\varepsilon = [\varepsilon; 1 - \varepsilon] \subset [0; 1], \quad (5.11)$$

where  $[\cdot]$  is the integer part operator. Without prior knowledge about the direction of a break, one is interested in testing the null against the two-sided alternative hypothesis  $H_A : \eta_{[nr]} \neq \eta$ . For practical reasons the above test is calculated over compact subsets of  $[0; 1]$ , i.e.,  $t$  equals the integer part of  $nr$  for  $r \in R_\varepsilon = [\varepsilon; 1 - \varepsilon]$  and for small  $\varepsilon > 0$ . Sets like  $R_\varepsilon$  are often used in the construction of parameter constancy tests (see, e.g., Andrews (1993)).<sup>11</sup> In line with Quandt's (1960) pioneering work on endogenous breakpoint determination in linear time

---

<sup>8</sup>Subscripts  $t$  now indicate that we relaxed the assumption of stationary tail behavior. All variables with  $t$  as subscript now refer to a subsample of the full sample  $1, \dots, n$ .

<sup>9</sup>Full sample values of  $m$  are determined by means of the Beirlant et al. (1999) exponential regression algorithm. In accordance with the minimization criterion of Goldie and Smith (1987), the theoretical value of  $m$  should be related to the sample size in a nonlinear way, i.e.,  $m = \kappa n^\gamma$ . Setting  $\gamma = 2/3$  and having obtained an estimate of  $m$  from the Beirlant algorithm we can solve for the scaling factor  $\kappa = m/n^{2/3}$ . Finally, subsample values for the recursive test can be determined using the scaling variable  $\kappa$ , i.e.,  $m_t = \kappa t^{2/3}$  with  $t$  the recursive subsample size.

<sup>10</sup>In case one uses this for the univariate return series one just has to replace  $Z$  by  $X$ .

<sup>11</sup>The restricted choice of  $r$  implies that  $\varepsilon n \leq t \leq (1 - \varepsilon)n$ . When the lower bound would be violated the recursive estimates might become too unstable and inefficient because of too small sub-sample sizes. On the other hand, the test will never find a break for  $t$  equal or very close to  $n$ , because the test value (5.10) is close to zero in that latter case. Thus, for computational efficiency one might stop calculating the tests beyond the upper bound of  $(1 - \varepsilon)n < n$ . We search for breaks in the  $[0.15n; 0.85n]$  subset of the total sample, as Andrews does.

series models, the candidate break date  $r$  can be selected as the maximum value of the test statistic (5.10), because at this point in time the constancy hypothesis is most likely to be violated.

Quintos et al. (2001) derived asymptotic critical values for the sup-value of (5.10) but these are not applicable in our framework. First, Quintos et al. (2001) assume that  $m$  is selected in such a way that the Hill estimator, stability test and resulting critical values are not marred by asymptotic bias. In practice, however, nearly all algorithms (including the Beirlant et al. (1999) algorithm that we implement) based on Asymptotic Mean Squared Error (AMSE) minimization induce an asymptotic bias term in the critical values. Also, the critical values can be further biased by nonlinear dependencies like, e.g., ARCH effects (volatility clustering).

We decided to determine the critical values by means of a parametric bootstrap of the recursive test while  $m$  and its subsample counterpart  $m_t$  are chosen by means of the Beirlant algorithm. In order to take account of the temporal dependence in the data and the possibility of volatility spillovers from one series to another, we use bivariate GARCH models as the basis for our parametric bootstrap. In order to keep the amount of parameters to be estimated as low as possible we chose for a diagonal BEKK(1,1,1) model first described by Engle and Kroner (1995). In general the BEKK(1,1,K) model can be defined as:

$$H_t = C'C + \sum_{k=1}^K A'_k \epsilon_{t-1} \epsilon'_{t-1} A_k + \sum_{k=1}^K G'_k H_{t-1} G_k, \quad (5.12)$$

where  $C$ ,  $A_k$ , and  $G_k$  are  $N \times N$  matrices,  $C$  is upper triangular and  $H_t$  is the conditional covariance matrix at time  $t$ . Thereby, the full model is being characterized by the following equation:

$$Y_t = \Gamma + \sigma_t, \quad (5.13)$$

where  $\sigma_t \sim N(0, H_t)$  and  $Y_t$  represents the a  $2 \times 1$  vector of the assets' returns at time  $t$  and  $\Gamma$  gives the average daily return. After estimating this model for all possible asset combinations we use the estimated coefficients and saved residuals for the parametric bootstrap of (5.10).

Quintos et al. (2001) report a Monte Carlo study that indicates good small sample power, size and bias properties of the recursive break test. Only in the case of a decrease of extreme tail dependence under the alternative hypothesis ( $\eta_1 > \eta_2$ ) they detect less acceptable power properties. We solve this problem by executing the recursive test both in a “forward” version and a “backward” version. The forward version calculates the sub-sample  $\eta$ s in calendar time, and the backward version in reverse calendar time. If a downward break in  $\eta$  occurs and the forward test does not pick it up, then the backward test corrects for this.

### 5.4.2 Cross-sectional variation

We would also like to know whether cross-sectional differences in linkage indicators for various asset pairs are statistically and economically significant. The asymptotic normality of  $\hat{\eta}$  enables some straightforward hypothesis testing.<sup>12</sup> However, equality tests based on the full sample values of the tail dependence parameter  $\eta$  are expected to be distorted if  $\eta$  values exhibit structural breaks. A test for the cross sectional equality of tail dependence parameters (null hypothesis) over time seems therefore more appropriate and can be based on the following statistic:

$$Q_t = \frac{\hat{\eta}_{1,t} - \hat{\eta}_{2,t}}{s.e.(\hat{\eta}_{1,t} - \hat{\eta}_{2,t})}, \quad (5.14)$$

with  $\hat{\eta}_{1,t}$  and  $\hat{\eta}_{2,t}$  standing for recursive estimates of the tail dependence of asset pairs to be compared. The test statistic should be close to normality provided  $t$  is sufficiently large.<sup>13</sup> Accordingly, the asymptotic critical values are 1.65, 1.96 and 2.58 for the 10%, 5% and 1% significance levels, respectively. In the empirical applications below the asymptotic standard error in the test's denominator (5.14) is estimated using a nonparametric asymptotic variance estimator proposed by Drees (2003) that is robust for general nonlinear temporal dependence in the data.

## 5.5 Extreme asset linkage results: Stocks, bonds, T-bills, gold

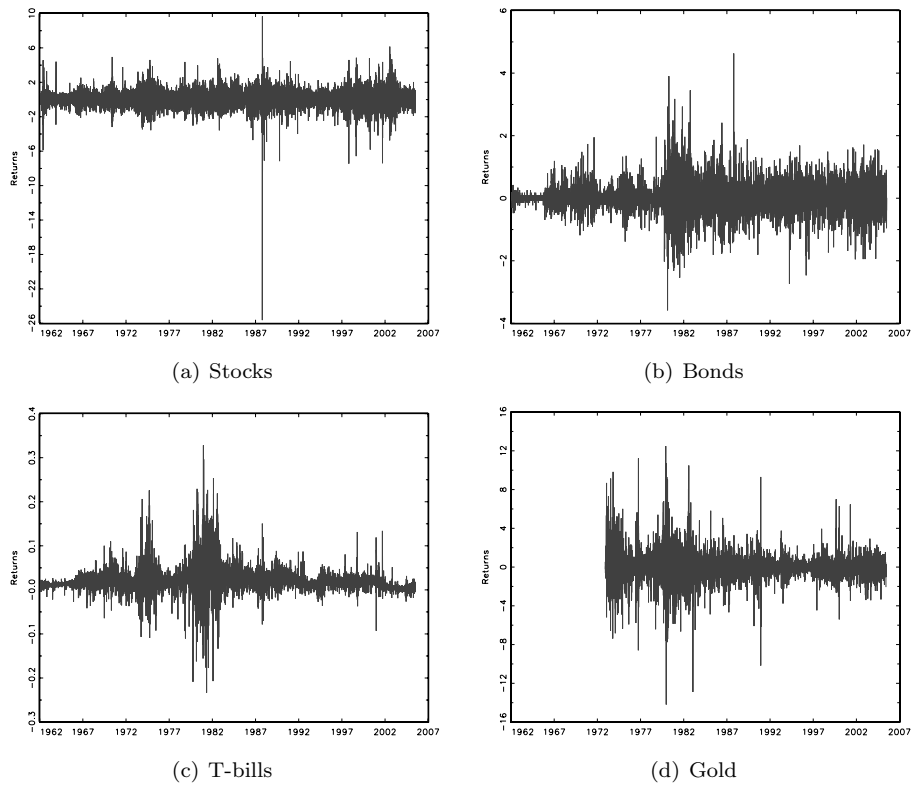
In this section we assess the likelihood of extreme return exceedances and co-exceedances for different asset classes in the U.S. financial markets. The data consist of 11,327 daily observations for stocks, bonds, and T-bills and 8,480 daily observations for gold. Time series for stocks, bonds and T-bills roughly span the period 1962-2005. We take the Dow Jones Industrials Index, ten year constant maturity government bonds, and three month constant maturity US government T-bills, respectively. Gold price series are significantly shorter and only start after the demise of the Bretton Woods system and the related abolishment of gold-US\$ convertibility in the beginning of the 1970s. A more detailed description of the data is given in the appendix. The series are plotted in Figure 5.1.

---

<sup>12</sup>Asymptotic normality of the estimator has been established in, for example, Hsing (1991), Quintos et al. (2001) and Drees (2003).

<sup>13</sup>One can safely assume that  $Q$  comes sufficiently close to normality for empirical sample sizes as the one used in this paper (see, e.g., Hall (1982), or Embrechts et al. (1997)).

Figure 5.1: Returns of stocks, bonds, T-bills, and gold



**Note:** Returns have been calculated as explained in the appendix.

### 5.5.1 Univariate results

In this section we analyze the tail behavior of univariate distributions of a sample of asset returns as a preliminary step for detecting possible extreme co-exceedances across asset classes. We also analyze the squared return tails and interpret it as a proxy of “extreme” volatility.

Descriptive statistics for all daily asset returns and squared returns are reported in Table 5.1. Only stock returns clearly exhibit negative skewness. Stocks and gold show the largest spread in their return distribution, which might be concluded from the maximum, minimum, and standard deviation measures. All series have excess kurtosis which indicates deviations from normality (fat tails).

Table 5.1: Descriptive statistics for daily US asset returns

Panel A: Returns								
Asset	Mean	Median	Max	Min	Std.Dev.	Skewness	Kurtosis	Obs.
Stocks	0.0237	0.0000	9.67	-25.63	0.9578	-1.73	53.33	11327
Bonds	0.0261	0.0227	4.63	-3.59	0.4285	0.18	10.03	11327
T-bills	0.0213	0.0186	0.32	-0.23	0.0244	1.49	21.81	11327
Gold	0.0221	0.0000	12.50	-14.20	1.3016	0.30	16.02	8480

Panel B: Squared Returns								
Asset	Mean	Median	Max	Min	Std.Dev.	Skewness	Kurtosis	Obs.
Stocks	0.00918	0.002260	6.570	0.0	0.06630	86.44	8497.66	11327
Bonds	0.00184	0.000290	0.210	0.0	0.00550	13.06	325.76	11327
T-bills	0.00001	0.000004	0.001	0.0	0.00003	12.17	240.79	11327
Gold	0.01690	0.002460	2.010	0.0	0.06570	13.40	268.82	8480

**Note:** The data are daily from the beginning of 1962 until the end of 2005. The observations for gold start in 1973.

### Univariate extreme value analysis

Table 5.2 summarizes the magnitude and timing of the two most extreme in-sample events together with the tail index and the percentile estimates based on equations (5.6) and (5.7), respectively. Panel A contains the results for the returns whereas the squared return results are reported in Panel B. Within Panel A we further distinguish between the left and right tail of the unconditional return distributions in order to account for possible asymmetries. Panel A shows that extreme losses and gains for stocks and gold are generally much higher than for bonds and T-bills. Even excluding the most extreme stock returns in October 1987 would not change this result. Moreover, for stocks and gold the historical extremes point toward tail asymmetries. The extreme negative returns are much larger in absolute value than the respective positive returns. For bonds and T-bills this is not so clear cut and tends to be the other way around.

Table 5.2: Minima, maxima, tail index, and univariate tail estimates for daily US asset returns

Panel A: Returns		Left Tail					Right Tail				
Asset		Min 1 (%)	Min 2 (%)	opt. m	$\hat{\alpha}$	percentile	Max 1 (%)	Max 2 (%)	opt. m	$\hat{\alpha}$	percentile
Stocks		-25.63	-8.38	287	3.42	$\frac{1}{10,000}$	9.67	6.15	189	3.69	$\frac{1}{10,000}$
		10/19/87	10/26/87			-9.24	10/21/87	07/24/02			8.76
Bonds		-3.59	-2.74	158	3.84	-4.01	4.63	3.90	145	3.68	4.37
		02/19/80	04/04/94			-2.20	10/20/87	04/16/80			2.34
T-bills		-0.23	-0.20	51	2.64	-0.27	0.32	0.29	781	2.64	0.61
		05/04/81	10/09/79			-0.11	12/19/80	01/05/81			0.25
Gold		-14.20	-12.89	103	3.12	-16.22	12.50	11.21	191	2.73	20.92
		01/22/80	02/28/83			-7.76	01/03/80	11/03/76			9.00

Panel B: Squared Returns		Left Tail					Right Tail				
Asset		Max 1 (%)	Max 2 (%)	opt. m	$\hat{\alpha}$	percentile	Test st.	P-value in %			
Stocks		6.57	0.93	579	1.74	$\frac{1}{10,000}$	-0.27	39.36			
		10/19/87	10/21/87			1.26					
Bonds		0.2143	0.1525	92	2.42	0.1621	0.10	45.84			
		10/20/87	04/16/80			0.063					
T-bills		0.0010	0.0008	56	2.67	0.0009	-0.00	50.07			
		12/19/80	01/05/81			0.0004					
Gold		2.0156	1.66	60	2.02	2.72	0.31	37.81			
		01/22/80	02/28/83			0.87					

Panel C: Test for tail index equality

**Note:**  $\hat{\alpha}$  is the reciprocal of the Hill estimator in equation (5.7). The columns "percentiles" are the percentiles with marginal probabilities of  $p = 1/10,000$  and  $p = 1/1,000$ , respectively. Max and Min 1 and 2 are the two most extreme positive and negative return observations in the sample, respectively. Panel A shows estimation results for the left and right tail. Panel B shows those for the squared returns as a proxy for volatility. Panel C gives test statistics and p-values for the test for equality of the left and right tail indexes as shown in (5.14).



A somewhat different picture emerges when we consider the estimated tail indices  $\hat{\alpha}$ . Stock returns seem to be asymmetric but the left tail index  $\hat{\alpha} = 3.42$  only slightly falls below its right tail counterpart ( $\hat{\alpha} = 3.69$ ). Bond and gold tail behavior suggest a fatter right tail. T-bills seem to exhibit symmetric tails. The left-tail index estimates are highest in the case of bond returns. Otherwise stated, long-term government-bond investments exhibit more limited downside risk than stocks. These results seem to confirm earlier research by e.g. Longin and Solnik (2001).

In Panel B we show the results for squared returns. Squared returns can be interpreted as a measure for assets' volatility. Engle (1982) pointed out that asset return volatility is likely to change over time but in a persistent manner. He developed a test for the so-called ARCH effect by choosing squared returns as a volatility proxy and regressing squared returns on lagged squared returns. It has been theoretically shown that there is a relation between volatility clustering and fat tails, see Koedijk and Schafgans (1973). Moreover, it can be shown that the squared returns should also be heavy tailed and that the probability mass in the tails of the return squares is even higher. Panel B reveals that the estimated tail indices for the squared returns are below the Panel A tail indices indeed.

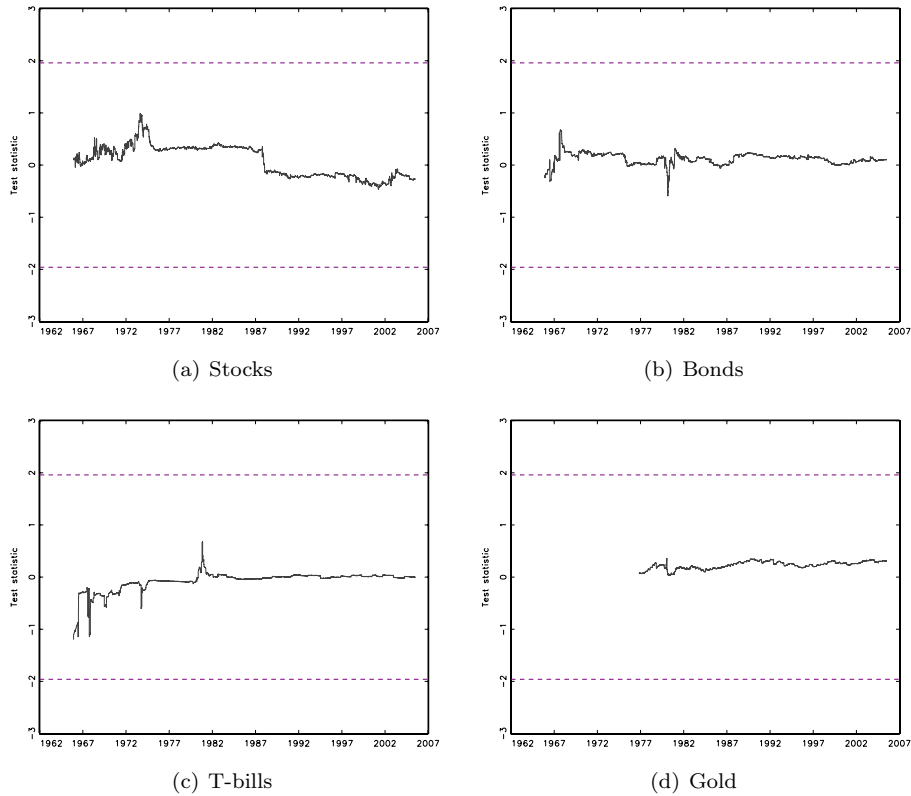
The table also provide some casual evidence for cross asset linkages during crisis periods. The calendar dates of the extreme events, as recorded below the minima and maxima, suggest the presence of a 'flight-to-quality' effect from stocks to bonds after Black-Monday. Stocks crashed on 10/19/1987 and bonds boomed on 10/20/1987. Notice also that the US stock market showed a strong technical upward correction on 10/21/1987 partly offsetting the exaggerated slump from two days before. Another interesting observation is that from the twelve most extreme events in the case of bonds, T-bills, and gold eight fall in the years between 1979 and 1981 which probably reflects that extreme volatility was at its highest around the second oil crisis. Similar results hold for the squared returns (volatilities). The most volatile period for stocks and bonds was in 1987. As for bonds, T-bills, and gold four out of the six most volatile days were in the period between 1979 and 1981.

The economic issue of interest, both for the general assessment of financial market stability and for financial investors' and institutions' risk management, is the likelihood and size of extreme returns as reflected by the tail probabilities and corresponding percentiles. The percentiles reflect possible extreme events or scenarios whose expected waiting time to occur equals the inverse of the corresponding marginal excess probabilities  $p$ . For example, a daily meltdown in the Dow Jones Industrial Average of -4.70% or more is expected to happen only once every 1,000 days or 3.9 years. So, the reported values can be interpreted as value-at-risk (VaR)

estimates for given marginal significance levels  $p$ .

The question remains whether the observed differences in tail index point estimates are statistically significant across tails. In order to test the null hypothesis of equal tail indices, we report corresponding test statistics for the tail asymmetry test and p-values (Panel C of the same Table 5.2). Additionally we show the recursive test version over the sample in Figure 5.2 for all four assets. Both the (full sample) test statistics in the table and the recursive statistics in the figure show that none of the assets exhibits significant tail asymmetry.

Figure 5.2: Recursive cross-section test: Univariate



**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels.

### Stability of EVT estimates

In order to check for structural breaks in the tail behavior of the unconditional return distribution we utilize a test developed by Quintos et al. (2001) as described in Section 5.4. Results of the test are summarized in Table 5.3. The table is again split into left and right tail results (except for the squared returns). In the most right part of the table we report the results for our volatility measure. Panel A states the results for the recursive test which checks for an increase in the thickness of the respective tail so a decrease in the tail index  $\alpha$ . In Panel B one can see the results for the reverse recursive test testing for the opposite. For every asset and tail<sup>14</sup> we report the test statistic, the bootstrap simulated critical value and the date of the break if the test is significant at least at the 5% level. Asterisks indicate the significance level of the test statistic. The only cases where the null hypothesis of tail index stability can be rejected at the 1% level involve the right tail of bonds and T-bills and the squared returns of bonds. Other identified breaks are significant at the 5% but not at the 1% level. Therefore, evidence for a structural break in the cases of the left tail for stocks and T-bills, the right tail for gold and the squared returns for T-bills and gold is much less convincing. In the case of bonds and T-bills one can clearly see that there appears to be a break in both series in the beginning of the 80s in the recursive test for the right tail. Later, in the mid and end 80s both right tails of the return distributions are detected to show again a break but now in the reverse recursive test. An obvious interpretation could be that Paul Volcker's structural change in the Fed's monetary policy and the second oil crisis had a major impact on the return behavior of bonds and T-bills but much less on stocks and gold. Volcker's shift from targeting the interest rates to rather limiting the growth rate of money supply had a strong influence on obligations' returns but also on their volatility. Thus, the turbulent times from the beginning of the 80s until black Monday in October 1987 are more strongly reflected in the return behavior of T-bills and bonds than in stocks and gold. Already a visual inspection of the return series in Figure 5.1 supports this result. As one can see, in the case of bonds the period starting in the early 80s until the end of the 80s shows stronger variability than the rest of the sample. For T-bills this unusual period is shorter, which is also reflected in the statistical test results. Somewhat surprisingly the test results only identify significant breaks for the right but not the left tail of the bonds and T-bills return distributions. One might attribute this to a relatively low power of the stability test.

---

<sup>14</sup>In the case of squared returns obviously only the right tail is being considered.

Table 5.3: Univariate results for stability test for daily US asset returns and squared returns

Panel A: Recursive Test

Asset	Left Tail			Right Tail			Squared Returns		
	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date
Stocks	6.12*	4.59	04/14/1986	1.39	3.58	-	1.74	5.15	-
Bonds	2.44	2.93	-	4.79**	2.11	02/18/81	4.08**	2.14	06/05/80
T-bills	1.20	2.24	-	24.83**	19.36	09/22/82	2.93*	2.87	05/19/80
Gold	0.53	3.68	-	0.27	2.12	-	0.21	2.77	-

Panel B: Reverse Recursive Test

Asset	Left Tail			Right Tail			Squared Returns		
	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date
Stocks	1.42	3.90	-	1.49	3.90	-	1.89	5.13	-
Bonds	1.63	3.02	-	3.43**	1.93	06/05/89	1.17	2.12	-
T-bills	3.43*	2.62	09/29/80	102.40**	19.36	07/08/85	4.18*	2.79	01/27/82
Gold	1.92	4.83	-	2.87*	2.35	06/14/00	3.24*	2.33	12/27/79

**Note:** Test statistics are based on equation (5.10). The bootstrap (BT) critical values were simulated as described in section 5.4 in the text. \* and \*\* indicate rejection of the null hypothesis of tail index constancy at the 5% and 1% significance levels, respectively.

### 5.5.2 Bivariate results

In this section we examine the propensity for co-exceedances across US asset classes. We start with standard correlation analysis followed by an identification of asset linkages during crisis periods using bivariate extreme value analysis. The extreme linkages allow us to assess the potential for cross asset substitution effects during market stress (the likelihood of flight to quality, flight to liquidity etc.).

#### Correlation analysis

The results of the correlation analysis are summarized in Figures 5.3 to Figure 5.8. For each of the six possible bivariate asset combinations we calculate rolling and recursive correlations. The rolling correlation is a yearly correlation in the sense that it is calculated using a time window of 260 trading days, which corresponds to one trading year. The recursive measure gives the correlation between the returns from the beginning of the sample period until point  $t$ . The correlation plots can be used to get some preliminary evidence for possible substitution effects between asset classes during periods of market stress.

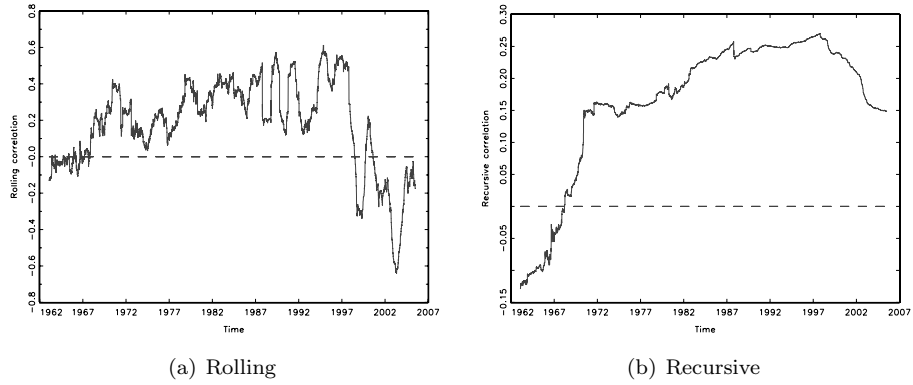
Correlations between stocks and bonds (Figure 5.3) or stocks and T-bills (Figure 5.4) are slightly more often positive than negative which does not seem to provide much evidence for substitution effects like flight to quality or flight to liquidity. However, the rolling correlations become negative around some crisis periods. For example, stock-bond (rolling) correlations turn negative after the Asian crisis and the negativity aggravates after the dotcom bubble burst. Stock-T-bill (rolling) correlations have similar signs around the same periods and also turn negative in the aftermath of the 1987 stock market crash. Thus, the rolling correlations provide some evidence of substitution from stocks into bonds or T-bills during crisis periods.

Correlations between stocks, bonds and gold more strongly point towards gold as a safe haven during times of market stress. First, the stock-gold (Figure 5.5) and bond-gold (Figure 5.7) rolling correlations tend to be more often negative than positive (both over crisis and noncrisis periods). Second, the negative correlations seem particularly present during the oil shocks in the 70s, the stock market crash of 1987, the 1997 Asian crisis and the 2000-2001 dotcom bubble burst.

Unsurprisingly, the correlations between bond and T-bill returns (Figure 5.6) are positive due to their linkage via the term structure. What the plots show is that this linkage can become weaker or stronger over time but it never becomes negative.

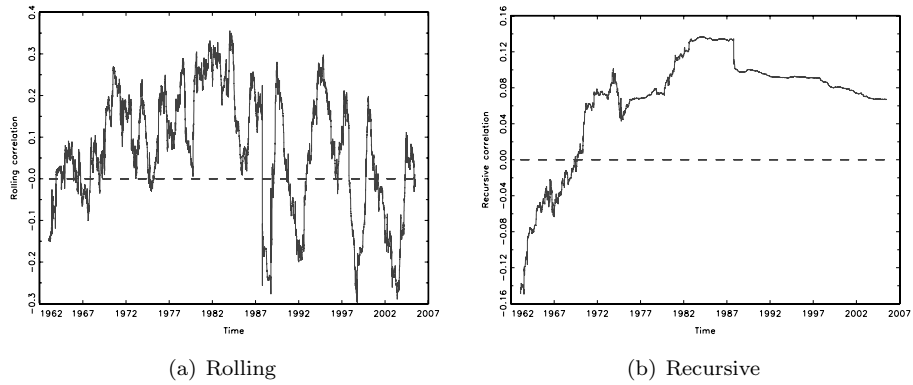
Finally, the correlations between T-bills and gold are not indicative of any

Figure 5.3: Stocks and bonds: Rolling and recursive correlation



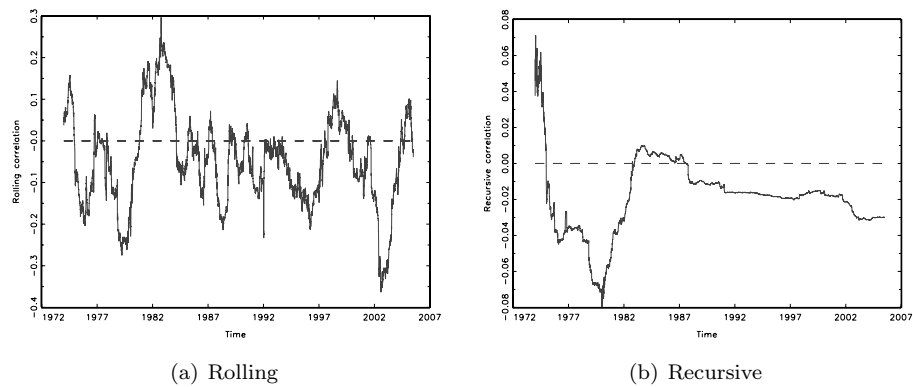
**Note:** Rolling correlations have been calculated with a window length of 260 trading days.

Figure 5.4: Stocks and T-bills: Rolling and recursive correlation



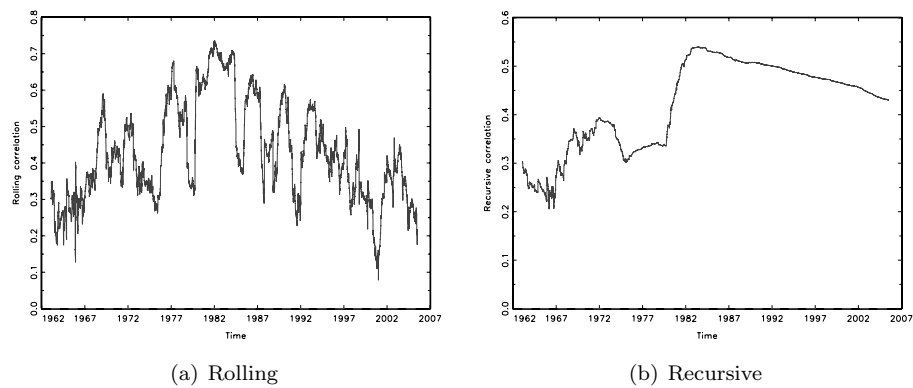
**Note:** Rolling correlations have been calculated with a window length of 260 trading days.

Figure 5.5: Stocks and gold: Rolling and recursive correlation



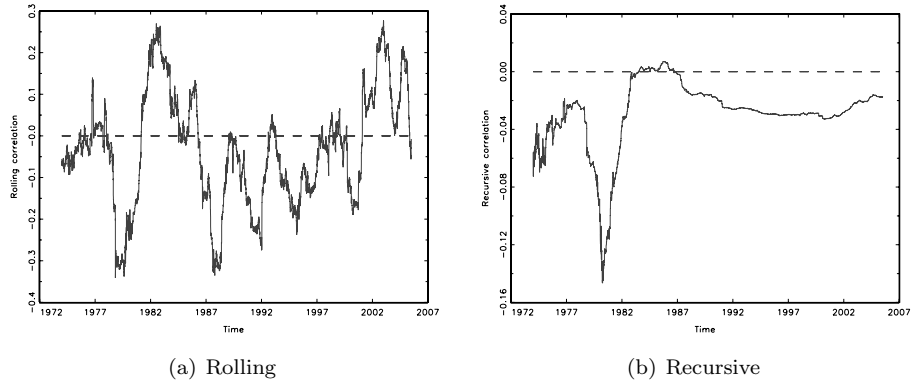
**Note:** Rolling correlations have been calculated with a window length of 260 trading days.

Figure 5.6: Bonds and T-bills: Rolling and recursive correlation



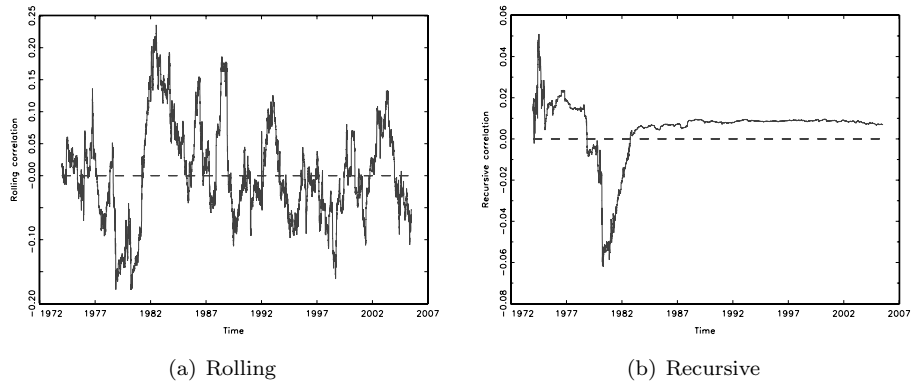
**Note:** Rolling correlations have been calculated with a window length of 260 trading days.

Figure 5.7: Bonds and gold: Rolling and recursive correlation



**Note:** Rolling correlations have been calculated with a window length of 260 trading days.

Figure 5.8: T-bills and gold: Rolling and recursive correlation



**Note:** Rolling correlations have been calculated with a window length of 260 trading days.



substitution effects. This should not surprise given that both assets are considered as interesting investment objects in times of distress.

There are numerous problems, however, with the use of correlations as dependence measures. First, they typically measure linear dependencies whereas it is often suggested that linkages during stress periods might be nonlinear phenomena. Otherwise stated, correlation figures not necessarily give a good indication for co-dependence of the extremes of the marginal return distributions. We therefore decided to apply a alternative framework that only exploits information for the bivariate tail.

### **Bivariate extreme value analysis**

Table 5.4 reports estimates for the conditional probability measure (5.1) and the conditional expectation measure (5.2) for all possible asset pairs in our sample assuming stationarity. The extreme measures are conditioned on different marginal excess probabilities  $p$  allowing us to evaluate the extreme dependence measures for different crisis levels.<sup>15</sup> Bivariate measures also allow us to compare the propensity towards co-crashes across assets with that towards substituting for a potentially safer asset (flight to quality or flight to liquidity effects).

The table reports the values for the tail index calculated in Equation (5.7), optimal amount of extremes  $m$ , conditional probabilities and expected values for occurrence of the mentioned co-exceedances corresponding to equation (5.1) and (5.2), respectively.<sup>16</sup>

Interpretation of the conditional probabilities is straightforward. For example the entry 2.33 for stock-bond co-crashes means that there is a 2.33% chance of a sharp joint drop in stock and bond values. “Sharp” in this context means that the crash levels correspond to 0.1% VaR for stock and bond tails. In Table 5.2 we saw that the univariate percentiles, on which we condition for calculating the stock-bond cocrash probability, correspond to -4.71% for stocks and -2.20% for bonds. The reader might be tempted to interpret the potential for stock and bond co-crashes as small. However, if the extreme events were independent, we would expect a conditional probability of around 0.1%. So, conditioning on a crash in one market, increases the probability that the other one also collapses by a factor 23. The co-exceedance probabilities in Table 5.4 all exceed the benchmark level of 0.1% implying that there is significant tail dependence. Stated otherwise, the probability of having an extreme gain or loss in one asset category suddenly becomes much higher once another “domino stone” has fallen.

---

<sup>15</sup>The lower the value of  $p$  the further we look into the bivariate tail.

<sup>16</sup>The optimal  $m$  is determined by means of the Beirlant et al. (1999) algorithm.

## 5.5 EXTREME ASSET LINKAGE RESULTS: STOCKS, BONDS, T-BILLS, GOLD

By further inspecting the table one can see that on average gold is less linked to the other assets during extreme events. The conditional probabilities are on average lower than in the other asset co-event cases. As such, gold looks as a reasonable hedge against all other asset classes considered.

Table 5.4: Bivariate results: Full sample

Panel A1: Stocks and Bonds, Stocks and T-bills, Stocks and Gold

	Stocks and Bonds		Stocks and T-bills		Stocks and Gold	
	Co-crash	FtQ S/B	Co-crash	FtL S/TB	Co-crash	FtQ S/G
Tail Index	1.45	1.622	1.52	1.84	1.82	1.79
Optimal m	290	404	397	424	371	437
Con.Pro. in %						
$p = 0.1$	2.33	0.68	1.54	0.23	0.29	0.37
$p = 0.05$	1.68	0.44	1.07	0.13	0.16	0.22
$p = 0.01$	0.79	0.16	0.47	0.03	0.04	0.06
E-Values						
$p = 0.1$	1.0118	1.0034	1.0082	1.0011	1.0014	1.0019
$p = 0.05$	1.0085	1.0022	1.0051	1.0009	1.0008	1.0011
$p = 0.01$	1.004	1.0008	1.0022	1.0004	1.0002	1.0003

Panel A2: Bonds and T-bills, Bonds and Gold, Gold and T-bills

	Bonds and T-bills		Bonds and Gold		Gold and T-bills	
	Co-crash	FtL B/TB	Co-crash	FtQ B/G	Co-crash	FtL G/TB
Tail Index	1.18	1.97	1.58	1.33	1.59	1.48
Optimal m	397	477	296	121	420	499
Con.Pro. in %						
$p = 0.1$	14.61	0.07	0.89	2.89	0.92	2.08
$p = 0.05$	12.83	0.03	0.61	2.31	0.61	1.51
$p = 0.01$	9.60	0.00	0.24	1.35	0.23	0.72
E-Values						
$p = 0.1$	1.0792	1.0003	1.0045	1.0146	1.0046	1.0105
$p = 0.05$	1.0691	1.0002	1.0031	1.0116	1.0031	1.0075
$p = 0.01$	1.0523	1.0000	1.0012	1.0068	1.0012	1.0035

**Note:** FtQ S/B stands for “flight to quality” from stocks into bonds, for example, as defined in the text. So, S stands for stocks, B for bonds, TB for T-bills and G for gold. E-values stands for the expected amount of co-events conditioned on an extreme percentile given by  $p$ .

When we have a look at the estimates for stocks and bonds, a co-crash is more likely than flight to quality from stocks into bonds. For stocks and T-bills we have lower estimates for the conditional probabilities and expected values, with flight to liquidity being close to the independence case. As for stock-T-bill co-crashes the probabilities are higher but still lower than for stock-bond co-crashes. So, capital leaving the stock market does not seem to cause a run on T-bills, i.e., the flight to liquidity hypothesis. Bonds and T-bills strongly co-move in the lower tails.

The bivariate results show that multivariate return distributions are not symmetric.

Notice that extreme event linkages can strongly differ from traditional dependence measures like correlation which is based on the full distributional support.

Otherwise stated, conclusions drawn from a simple full sample correlation or even a rolling correlation analysis can be misleading for the occurrence probability of extreme co-events. In order to show this, we can again refer to the correlation Figures 5.3 to 5.8. The full sample correlation is equal to the last observation of the recursive correlation. In the case of bonds and T-bills the full sample correlation is equal to 43.02%. One can interpret such a correlation as a strong linkage, but this can be rather illusory. If one uses the bivariate normal distribution in order to assess the extreme bonds-T-bills market linkage applying the sample variances and correlation, one would find a conditional co-crash probability of around 0.000125% for the marginal distribution percentile  $p = 1/1000$ . Using the more accurate EVT estimation of tail behavior we find for the same marginal percentile a conditional co-crash probability of 14.61%. Hence, correlations together with the multivariate normality assumption strongly understate extreme financial market linkages and thereby should not be used for extreme dependence estimation.

### **Stability of EVT estimates**

Here we relax the stationarity assumption of the dependency measures presented in the section before. Results are summarized in Table 5.5. Panel A1 and B1 show the results for the recursive test (alternative hypothesis = increase in tail dependence) and Panel A2 and B2 give the results for the backward recursive test (alternative hypothesis = decrease in tail dependence).

In Table 5.5 a clear picture emerges. The recursive test tends to find breaks in the early part of the sample ranging from 1968 until 1980 depending on the asset combination. The only exceptions are the asset combinations stocks with gold and gold with T-bills. A likely explanation is that the gold time series only starts in 1973 so that the break testing procedure only starts in the year 1976. But at that point the break might already have occurred such that the test is unable to pick it up.

Another interesting observation is that in all but one asset combination (stocks and bonds) we find significant breaks for the reverse recursive test all happening after the respective breakpoints for the recursive test. The breaks in the reverse recursive procedure range from 1983 until 1991 and most of them cluster between 1983 and 1988. So, considered market comovements tend to become stronger in the seventies and again weaker in the eighties. In most of the cases where the recursive test detects a break, the tail index becomes smaller either shortly before the first oil shock in the beginning of the 70s or before the second oil shock and the beginning of Volcker's term at the end of the 70s. In sum, the economically and politically turbulent times surrounding the oil shocks, the Volcker presidency of the FED and extreme market volatility around Black Monday in October 1987

## 5.5 EXTREME ASSET LINKAGE RESULTS: STOCKS, BONDS, T-BILLS, GOLD

Table 5.5: Bivariate results: Stability test

Panel A1: Recursive test

	Stocks and Bonds		Stocks and T-bills		Stocks and Gold	
	Co-crash	FtQ S/B	Co-crash	FtL S/TB	Co-crash	FtQ S/G
1969	22.3**		28.4**			
	09/25		06/10			
1973				25.35**		
				01/26		
1979		46.0**				
		02/22				

Panel A2: Backward recursive test

	Stocks and Bonds		Stocks and T-bills		Stocks and Gold	
	Co-crash	FtQ S/B	Co-crash	FtL S/TB	Co-crash	FtQ S/G
1983						8.5*
						03/11
1988				13.3**	10.0**	
				05/24/88	02/04/88	
1989			42.7**			
			04/13/89			

Panel B1: Recursive test

	Bonds and T-bills		Bonds and Gold		Gold and T-bills	
	Co-crash	FtL B/TB	Co-crash	FtQ B/G	Co-crash	FtL G/TB
1968	38.97**					
	11/29					
1973		55.5**				
		12/05				
1979				8.8**		
				09/28		
1980			17.1**			
			01/16			

Panel B2: Backward recursive test

	Bonds and T-bills		Bonds and Gold		Gold and T-bills	
	Co-crash	FtL B/TB	Co-crash	FtQ B/G	Co-crash	FtL G/TB
1983	54.6**				50.2**	
	10/12				11/24	
1986			42.9**			
			10/15			
1987				46.0**		
				12/29		
1991		45.8**				408.8**
		06/07				05/30

**Note:** FtQ S/B stands for “flight to quality” from stocks into bonds, for example, as defined in the text. So, S stands for stocks, B for bonds, TB for T-bills and G for gold. \* and \*\* indicate rejection of the null hypothesis of tail index constancy at the 5% and 1% significance levels, respectively.

seem to coincide with breaks in the degree of tail dependence. Moreover, most of the asset co-event cases show a tail index behavior that might be described as a U-shape or mean-reverting. As such, the full sample results in Table 5.4 represent an average across time. Nevertheless, those calculated conditional probabilities provide a good approximation of the true co-dependencies of assets because of the observed mean reverting behavior of the tail indices. A further interesting step could be to split the sample at the observed break points and estimate co-dependencies within every sub-sample. A problem of such an approach, though, constitutes the exact location of the break in the bivariate distributions, because the breaks across co-boom, co-crash, and flight to quality co-events do not always occur at the same point in time. In order to account for this, some kind of common break point estimation is needed but this is beyond the aim of this chapter.

Table 5.6 shows that the results for the squared returns are pretty similar to those of the ordinary returns. Stocks and bonds and especially bonds and T-bills show the highest conditional probability of common extreme volatilities. In the case of stocks and bonds, for example, we estimate a conditional probability of 5.64% that bonds show a volatility among their 0.1% largest ones, given that also stocks' volatility was among their 0.1% largest daily observed volatilities. This constitutes a probability increase by a factor 56 compared to the independence case. For the asset combination bonds and T-bills the same estimated conditional probability even increases with a factor 232. For the other four possible bivariate asset combinations, estimated conditional probabilities for pairs of squared returns are much lower also supporting the results obtained before.

Table 5.6: Bivariate results: Squared returns

	S & B	S & TB	S & G	B & TB	B & G	G & TB
Tail Index	1.29	1.59	1.74	1.04	1.47	1.65
Optimal m	330	498	462	443	426	499
Con.Pro. in %						
$p = 0.1$	5.64	1.10	0.5	23.20	2.16	2.99
$p = 0.05$	4.61	0.73	0.31	22.52	1.56	2.19
$p = 0.01$	2.86	0.28	0.11	21.01	0.74	1.05
E-Values						
$p = 0.1$	1.029	1.006	1.003	1.131	1.011	1.015
$p = 0.05$	1.024	1.004	1.002	1.127	1.008	1.011
$p = 0.01$	1.015	1.001	1.001	1.117	1.004	1.005

**Note:** S stands for stocks, B for bonds, TB for T-bills and G for gold. E-values stands for the expected amount of co-events conditioned on an extreme percentile given by  $p$ .

The structural break point analysis for the squared returns in Table 5.7 is generally in line with the results of the ordinary returns. Probabilities of common high volatility tend to increase either before the first or the second oil crises and

decrease again at the end of the 80s and the beginning of the 90s. So, on average volatilities' tail indexes tend to co-break with the ordinary returns.

Table 5.7: Bivariate results: Stability test squared returns

Panel A1: Recursive test

	S & B	S & TB	S &	B & TB	B & G	G & TB
1969	22.93	41.84				
	10/08	02/19				
1974				76.76		
				07/19		
1978					20.89	
					08/10	

Panel A2: Backward recursive test

	S & B	S & TB	S &	B & TB	B & G	G & TB
1988			11.84		76.52	
			02/04		02/01	
1989				126.12		
				06/26		
1990		28.54				
		10/12				
1991						455.39
						02/28

**Note:** S stands for stocks, B for bonds, TB for T-bills and G for gold.  
All breaks are found to be significant at a 5% significance level.

A little word of caution might be appropriate here. In the transformation to the Pareto distribution we implicitly assume stationarity of the univariate return series' tail behavior. We know from Section 5.5.1 that this is not always the case. Nevertheless, we believe that following our approach we are able to distinguish the cases of having only a break in the marginal return distributions or having a break in dependence structure of the marginal distributions. This can also be confirmed by comparing the univariate and bivariate break dates, which do not coincide. If our approach was not able to distinguish both cases break dates would have to coincide. Future research on the theoretical foundations would be very interesting but goes beyond the scope of this chapter.

### Cross-sectional results

In this subsection we apply the same cross section test for comparing bivariate tail indices across assets as we did for comparison of the univariate indices in Table 5.2 and in Figure 5.2. Here, we are interested if the co-exceedance tail indices (tail dependence parameters) differ across asset combinations and between co-crashes and/or flight to quality/liquidity. Smaller tail indices here mean that the corresponding co-extreme events are more likely to occur than one with a bigger tail index. In Figures 5.9 to 5.15 we show the recursive test statistics as in

Equation (5.14). Figures 5.9 to 5.11 show all co-crash combinations, Figures 5.12 to 5.14 are for all flight to quality/liquidity combinations and Figure 5.15 gives the test for all matched cases of co-crashes and flight to quality/liquidity. The figures also show the upper and lower rejection regions at -1.96 and 1.96 corresponding the normal distribution 2.5% levels each. Negative test statistics indicate a smaller tail index for the first pair of assets compared to the second pair. Two examples make the logic clear. In Figure 5.9 (a) we have the case of the co-crash combination of stocks and bonds with stocks and T-bills. So, we actually compare two bivariate time series' tail indices. In this specific case there does not appear any significant difference between both tail indices through the full sample period. As a second example serves Figure 5.15 (a) which corresponds to the case where we compare the  $\alpha$  of the bivariate series where bonds and T-bills crash together, against the case where we speak of flight to liquidity from bonds to T-bills. Here we see that for most of the time series the tail index for bonds and T-bills co-crashes is significantly smaller (positive test statistics) than the tail index for the flight to liquidity case from bonds into T-bills.

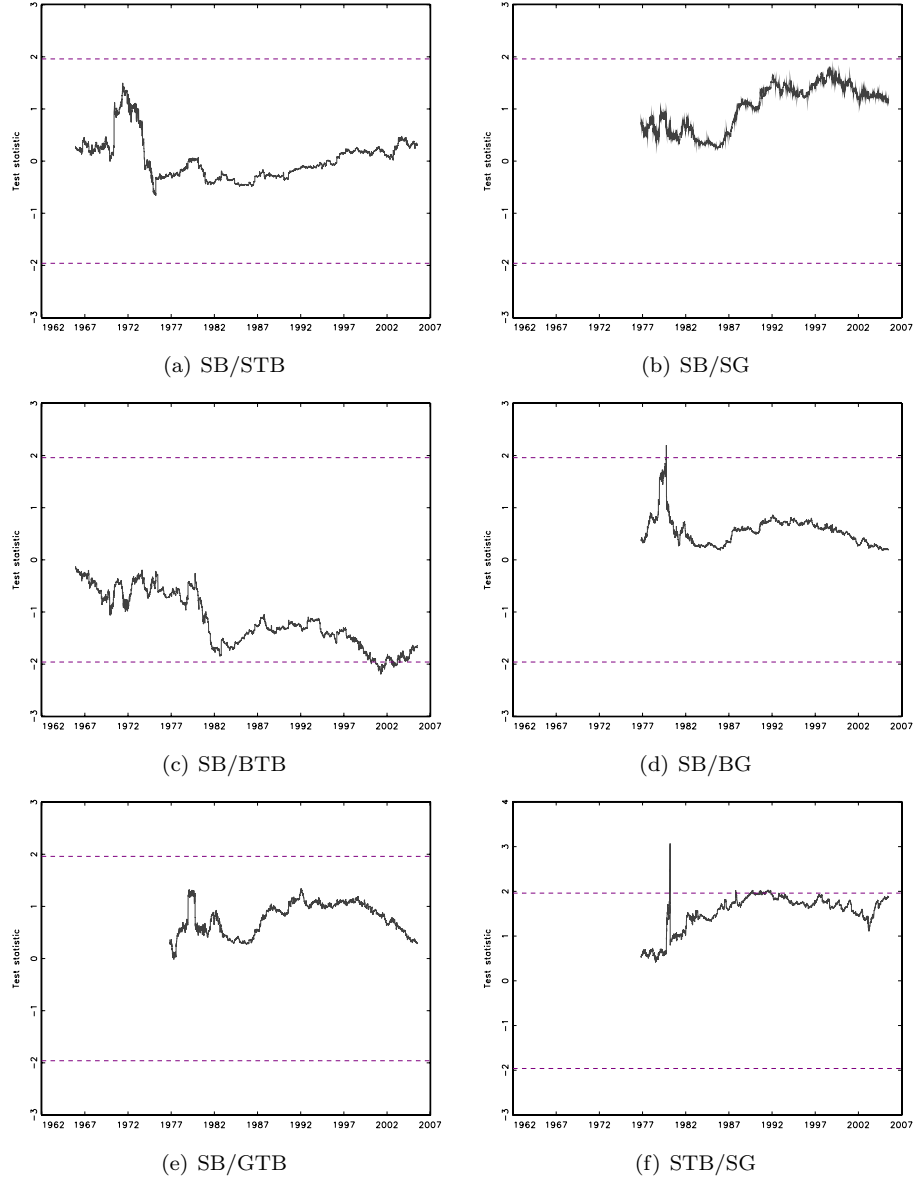
We can draw a couple of interesting conclusions from those figures. First, there are only a few significant differences between bivariate tail indices over the full sample period and all asset combinations. This confirms the general findings in the EVT literature that tail indices usually cannot be found to differ significantly.<sup>17</sup> Second, those cases where we do clearly find significant results always include the asset pair bonds and T-bills. Actually, this was to be expected. The pair bonds and T-bills shows clearly the strongest comoving behavior. Along with this, the probability that bonds crash and T-bills boom (flight to liquidity) at the same time is very unlikely. So, asset combinations including bonds and T-bills will have the tendency towards significant differences in the tail indices, which is what we find. Third, the cases involving bonds and T-bills also seem to be the most volatile in terms of movements of the test statistic. Although results have to be interpreted with caution, a possible explanation is that the bonds and T-bills returns are heavily influenced by changes in the monetary regime. Changes in the Fed's policy will directly move those securities' prices and thereby possibly lead to more changes in their comovement tail indices.

As one good example can serve Figure 5.10 (a) where we compare the tail indices for co-crashes between stocks and T-bills and bonds and T-bills. Here we see in the 1980s a period with a significantly smaller tail index for bonds and T-bills co-crashes than for stocks and T-bills co-crashes. Again this might be explained by the fact that Paul Volcker was the Fed's chairman and especially bonds markets

---

<sup>17</sup>This confirms earlier findings like in Koedijk and Schafgans (1990), Jansen and de Vries (1991) Quintos et al. (2001).

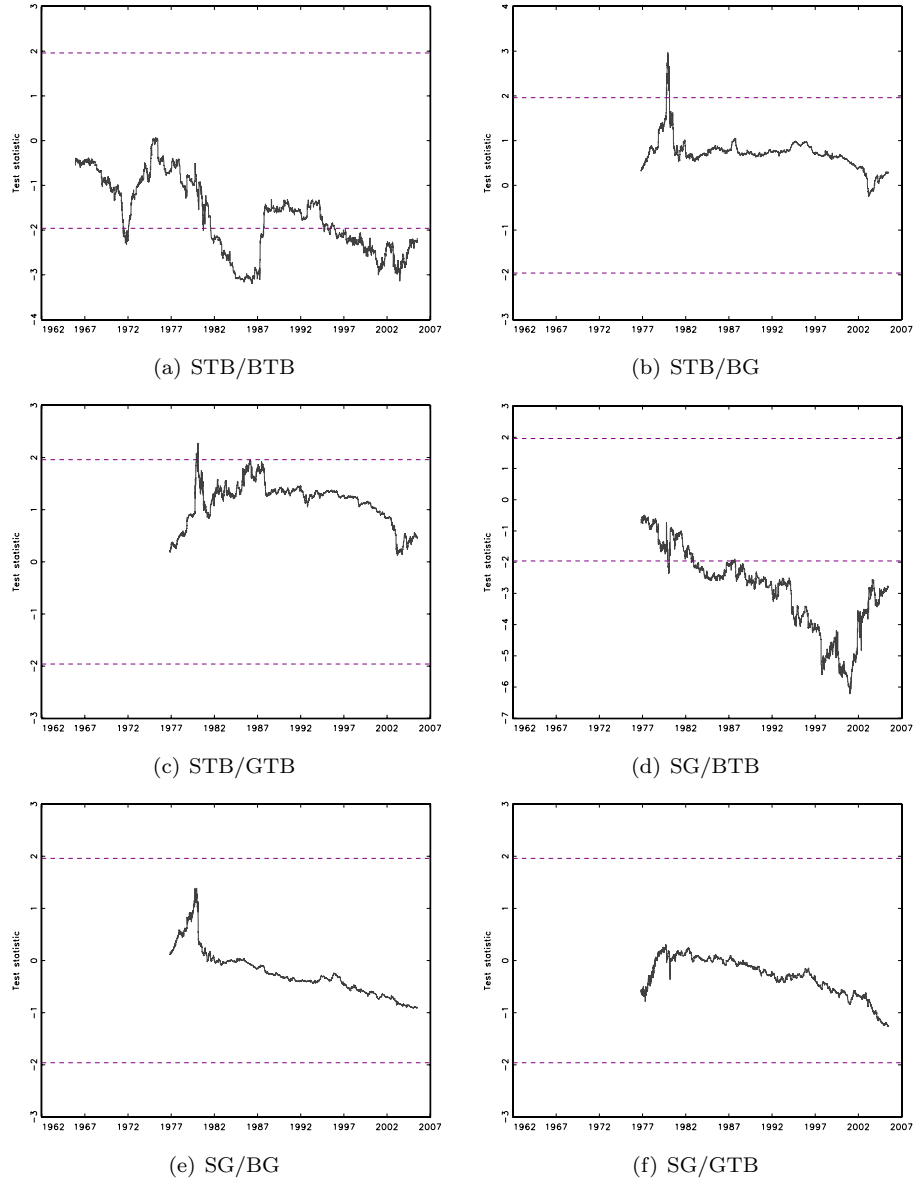
Figure 5.9: Recursive cross-section test: Co-crashes I



**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

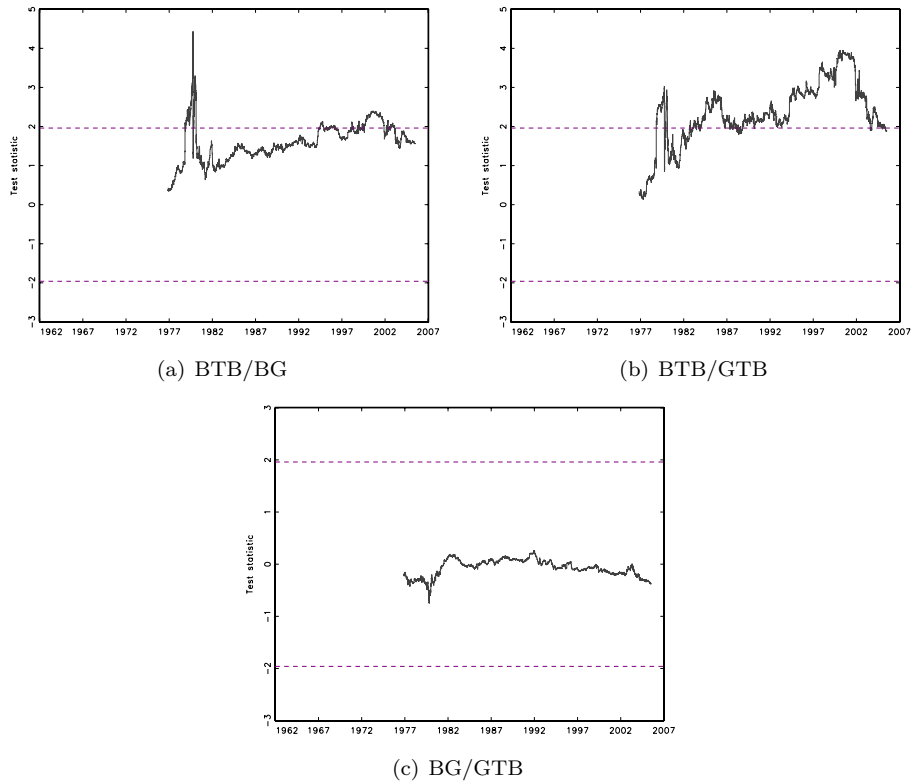


Figure 5.10: Recursive cross-section test: Co-crashes II



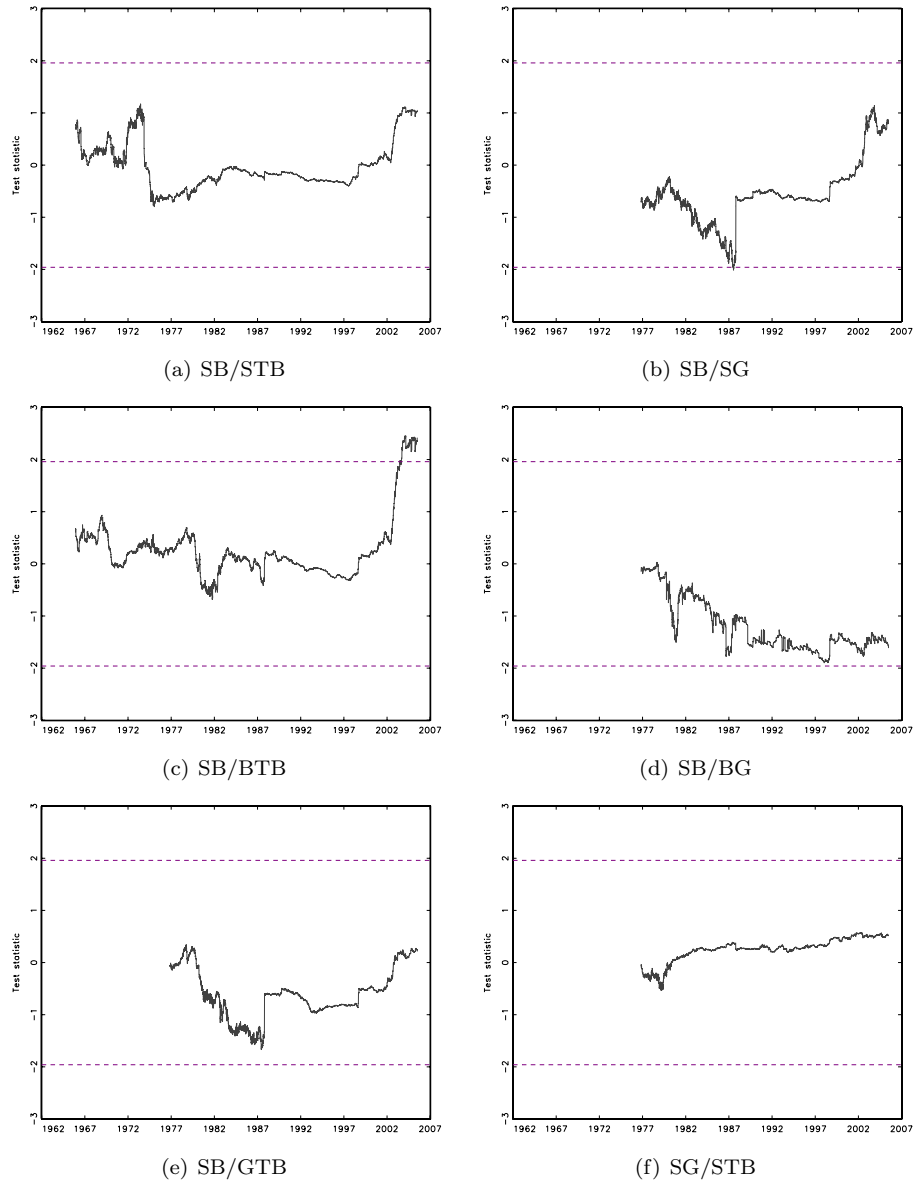
**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 5.11: Recursive cross-section test: Co-crashes III



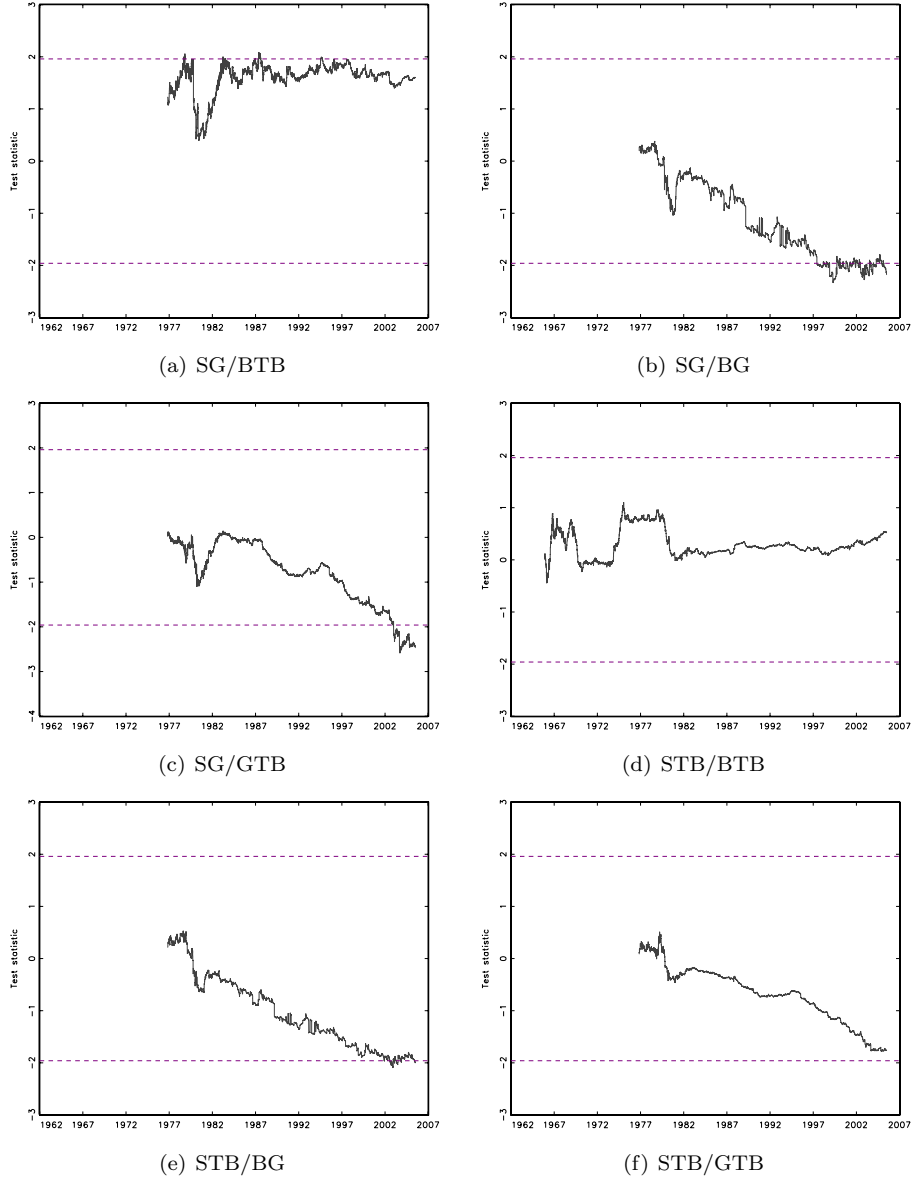
**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 5.12: Recursive cross-section test: Flight to quality and liquidity I



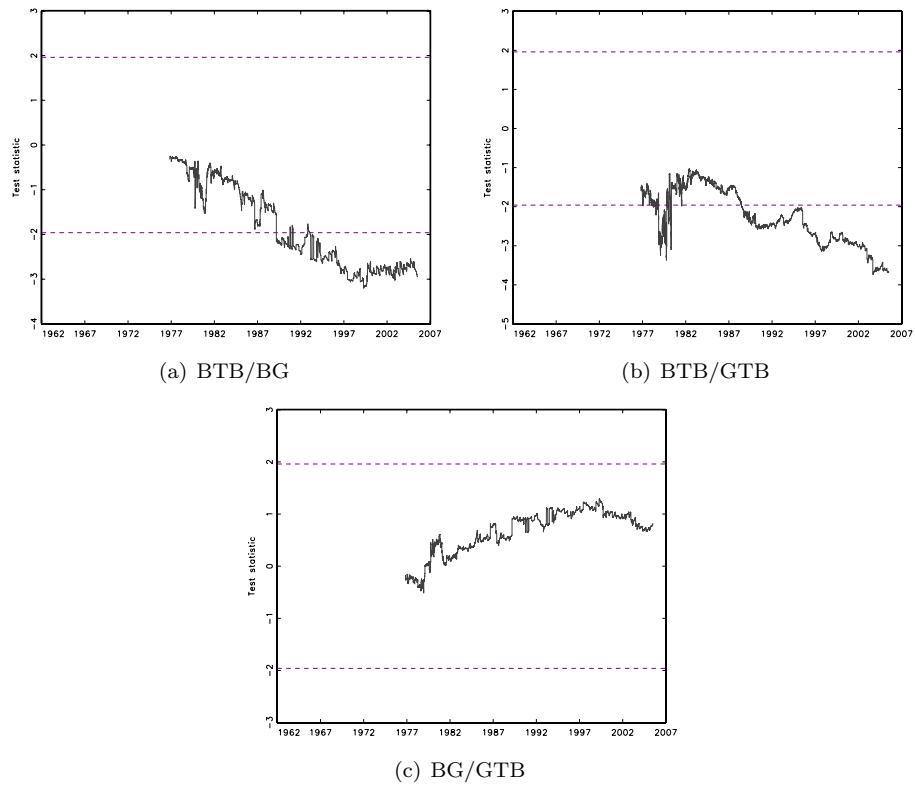
**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 5.13: Recursive cross-section test: Flight to quality and liquidity II



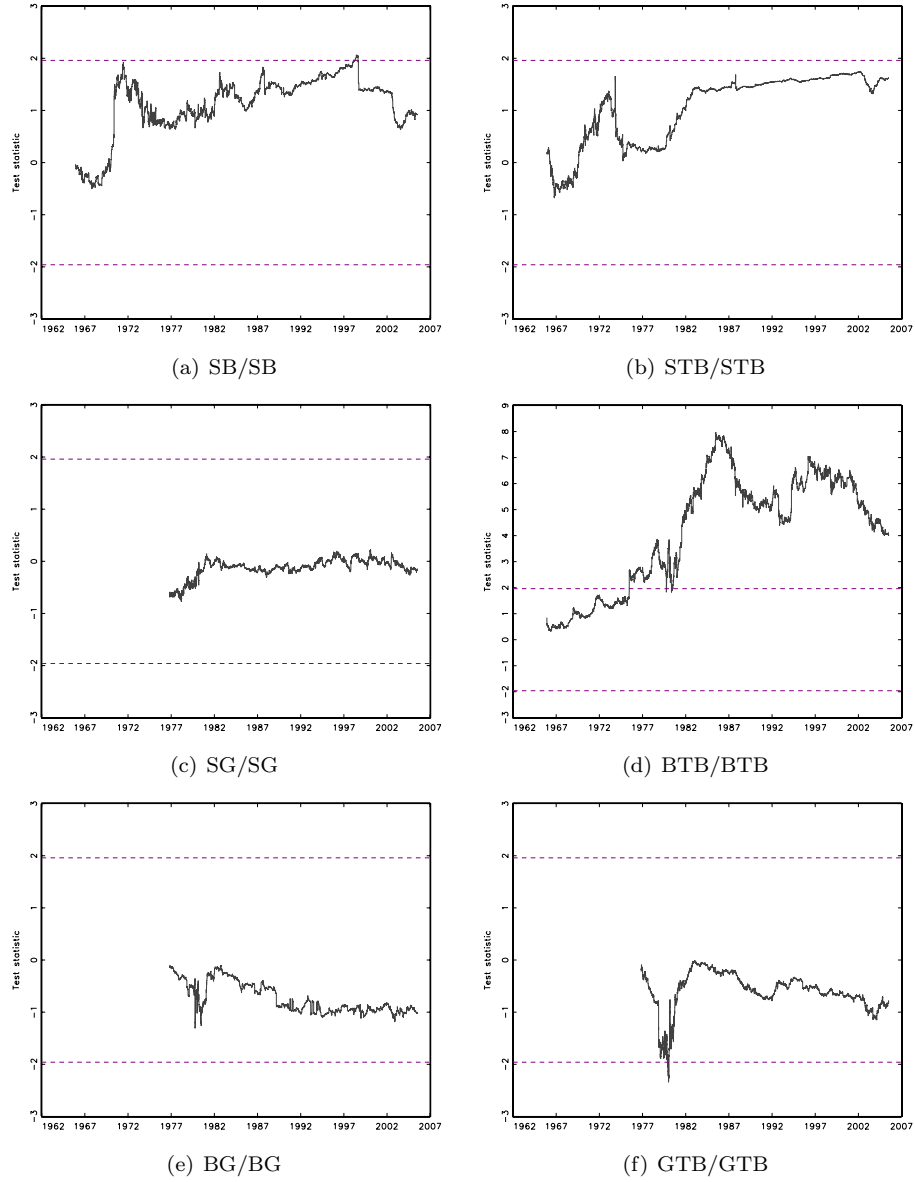
**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 5.14: Recursive cross-section test: Flight to quality and liquidity III



**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 5.15: Recursive cross-section test: Co-crashes and flight to quality/liquidity



**Note:** Test statistics have been calculated using Eq. 5.14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

were characterized by high volatility. In the year 1987 there is a strong change toward insignificance, probably caused by the Black Monday stock market crash and following volatility clearly having an impact on the lower tail index of stocks in that period.

## 5.6 Conclusions

In this chapter we study the linkages between four different US asset classes (US stocks, government bonds, T-bills, and gold) in times of market turbulence. The linkages were characterized by their asymptotic tail dependence. Studying the likelihood of co-exceedances across asset classes is interesting for investors who have to choose their investment portfolios according to their preferred risk-return combinations. On the other hand, policy makers and supervisory bodies are interested in these linkages because they potentially influence the level of systemic risk in financial markets.

We use a non-parametric multivariate measure to identify the tail dependence of the marginal return distributions and derive estimates for the expected conditional probabilities of return co-exceedances. Such an approach does not rely on a particular probability law for the marginal return distributions and thereby has advantages over the usual conditional correlation measures because wrong parametric assumptions can easily distort extreme-spillover probabilities. We also tested for stability of these linkage measures through time in order to see if there are any periods with stronger or weaker spillover effects.

A preliminary (rolling and recursive) correlation analysis showed similar positive correlations between stocks and bonds as well as stocks and T-bills over time. During periods of market turbulence (oil crises in the 1970s and 1980s, the stock market crash in October 1987 and the terrorist attacks in New York in 2001), correlations typically dropped to below-average values. As expected, T-bills and bonds show a very strong correlation (but decreasing since the second oil shock). Gold turns out to be only mildly correlated with the rest of the assets in the sample set. These time varying correlations already give a first indication that there are possible flight-to-quality and flight-to-liquidity effects (especially during periods of financial uncertainty).

Structural break tests applied on the univariate and multivariate extreme value measures show very similar results. Both cases reveal breaks from thin to fatter tails either before the first or the second oil shock. Shifts back from fat to thinner tails occur during the 1980s. This indicates that asset returns (especially stocks, bonds, and T-bills) showed an increased probability of co-exceedances in the period between the mid 70s and mid 80s. For the squared returns, results are rather

similar. Stocks and bonds, and especially bonds and T-bills show the highest conditional probability of common extreme volatilities. Again, breaks seem to occur around economically and politically turbulent times like the two oil shocks, the Volcker FED-presidency, or the stock market crash of 1987.



## 5.7 Appendix

We choose the Dow Jones Industrial Average as stock price index and extracted it from Datastream, Inc. 10-year government bond and 3-month T-bill returns were calculated from the corresponding yield to maturity data from the Board of Governors of the Federal Reserve Board.<sup>18</sup> We calculated returns from these yield data according to the methods used, for example, in Campbell et al. (1997, Chapter 10). Gold prices were extracted from [www.usagold.com](http://www.usagold.com). The stock data are Financial Times Standard & Poors world price indices, whereas the bond data correspond to ten year (“all-traded”) government bonds. In order to arrive at the returns for stocks and gold we calculated the first differences of the log of their prices levels. Data on stocks, T-bills, and bonds start on February 2, 1962 whereas gold starts on January 2, 1973. Last observations in the sample are on July 5, 2005. We did not include corporate bond indices, because of our particular interest in the flight to quality phenomenon. The stock and bond returns are not compensated for dividends and coupon payments, respectively.

---

<sup>18</sup><http://www.federalreserve.gov>.

---

# ESTIMATING AND FORECASTING ASSET VOLATILITY AND ITS VOLATILITY: A MARKOV-SWITCHING RANGE MODEL

In the following we propose a new model to describe asset market volatilities. It suggests to use the range, defined as the percentage difference between the maximum and the minimum price of an asset, as a nonparametric proxy for an asset's volatility. We find that range time series can very well be modeled in a Markov-switching (MS) autoregressive moving-average (ARMA) with generalized autoregressive conditional heteroscedasticity (GARCH) framework. Such a model allows to identify different volatility regimes, i.e. a high and a low volatility regime, and to forecast volatility  $h$  periods into the future. In an international comparison we find evidence for volatility comovements across different countries.

## 6.1 Introduction

Asset-volatility is of outstanding importance in finance. It directly and indirectly influences asset pricing (for example options prices directly depend on the underlying asset's volatility), the optimal hedge ratio, portfolio decompositions, and risk management among others (Alizadeh, Brandt, and Diebold, 2002). Therefore, volatility-modeling has been a focus of much academic research in the last decennia. Early contributions assumed constant asset-volatility (e.g., Merton (1969) or Black and Scholes (1973)). Especially the work of Engle (1982) and Bollerslev (1986), however, contributed to the nowadays widely accepted conviction that volatility is both time-varying and predictable (see also Andersen and Bollerslev (1997)). Engle introduced the autoregressive conditional heteroscedasticity

(ARCH) models whereas Bollerslev extended those to the class of generalized ARCH (GARCH) models. The observation that some time periods are affected by very high while others by relatively low volatility fostered the more recent development of regime-switching models. Building on Hamiltons (1989) work, Hamilton and Susmel (1994), Gray (1996) and Klaassen (2002) further introduced regime-switching ARCH and GARCH models improving further the modeling of volatility. ARCH and GARCH models are today the workhorse of asset-volatility modeling both in academics and industry (see for example Ghysels et al. (2006)).

Volatility in economics is defined as the variability of a random variable of a time series. Hence, this volatility is “...inherently unobservable, or latent, and develops stochastically through time” (Brandt and Diebold, 2006, p. 1). Volatility is inherently latent because the true data generating process of asset prices is not known, making it impossible to quantify unambiguously the “random component” of a time series and even more difficult to pin down its instantaneous variability or volatility. There appear to be two solutions to this problem. First, one can try to model the latent variable volatility as the conditional second moment/variance of an observed return series parametrically (e.g., Engle (1982), Taylor (1982), and Bollerslev (1986)). Or second, one uses nonparametric estimators for the volatility. The range, defined as the difference between the maximum and the minimum log asset prices over a fixed interval, appears here as a natural estimator and has indeed been the subject of much academic research (e.g., Garman and Klass (1980), Parkinson (1980), Beckers (1983), Ball and Torous (1984), Rogers and Satchell (1991), Andersen and Bollerslev (1998), Yang and Zhang (2000), Alizadeh et al. (2002), Chou (2005), Brandt and Diebold (2006)). Another interesting and related approach is introduced in Cheung (2007) where the daily highs and lows of the price process are modeled by means of cointegration analysis.

In contrast to the conditional variance modeling, the range is directly observable from the data and does not need to be estimated. Apart from being a very intuitive and directly observable volatility estimator, the range also is very efficient. Indeed, as noted in Brandt and Diebold (2006, pp.61):

As emphasized most recently by Alizadeh et al. (2002), the range is a highly efficient volatility proxy, distilling volatility information from the entire intraday price path, in contrast to volatility proxies based on the daily return, such as the daily squared return, which use only the opening and closing prices.

Moreover, as has also been mentioned by many authors (e.g., Alizadeh et al. (2002)), range data on the one hand are available for many different assets such as individual stocks, stock indices, currencies, and Treasury securities, and on the other hand these data series often have a history of many decades. This consti-

tutes a strong advantage over another nonparametric estimator of the variance, namely the realized volatility, which uses high-frequency data at say 5-minute intervals. Those data often only start in the middle of the 1990s if available at all. Furthermore, Alizadeh et al. (2002) clearly show that market microstructure problems like bid-ask spreads can severely bias the realized volatility estimator. Primarily theoretical references regarding realized volatility include, among others, Barndorff-Nielsen and Shepard (2001; 2003; 2004), and Andersen, Bollerslev, Diebold, and Labys (2003).

A further advantage of using an observed volatility estimator is that it can be modeled in the mean equation. This enables the econometrician to model the volatility of the volatility as the conditional second moment of the range in contrast to having to model it as the conditional fourth moment of a return series. Modeling the volatility of the volatility is important, for example, in option pricing, where an option trader wants to know the probability that the volatility, a direct price determinant, changes in order to optimize his pricing decision. Additionally, changes in volatility also have an influence on optimal hedge ratios (e.g., Ederington (1979), Lien and Tse (2002)). Therefore, predictable volatility of volatility can help in making better hedging decisions.

In the literature asset markets have been found to show regime-switching behavior. There are clear periods of low/normal volatility but also longer-lasting periods where asset market volatility is significantly higher than in the low-volatility periods. Such regime-switching volatility behavior has usually been modeled with first-order Markov processes. See, for example, Hamilton and Susmel (1994), Gray (1996), and Klaassen (2002).

Motivated by these points, we propose a simple yet efficient way of modeling asset market volatility and its volatility. We suggest to fit the log range of assets to a Markov-switching-(MS-)ARMA-GARCH time series model. This combines the advantages of the range as a nonparametric yet highly efficient volatility estimator with well established time-series modeling techniques in order to estimate and forecast asset volatilities. We fit our proposed model to weekly S&P500, ten year T-notes, three months T-bills, FTSE100, and Nikkei225 data. Our findings are: First, our model is well able to distinguish low from high volatility regimes. Second, that volatility dynamics change with the regime, which has important effects for forecasting purposes, confirming results found in Gray (1996). Third, a forecasting exercise leads to promising results by showing that some specifications of the model are able to clearly decrease forecasting errors with respect to a linear model.

This chapter proceeds as follows. After presenting the theoretical background about the range in Section 6.2 we describe the methodology for model estimations

in Section 6.3. Section 6.4 presents the results of the application of our model to the data of different assets and markets. We also present results of a forecasting comparison. Finally, Section 6.5 summarizes our results, concludes and sketches directions for future research.

## 6.2 Theoretical foundations

In this section we will present the theoretical distributional foundation of the range as an estimator for the diffusion constant of a continuous random walk that has already been derived in the literature. Next we heuristically argue that those distributional results should basically remain unchallenged once we allow for non i.i.d. errors. We also conduct a short Monte Carlo experiment to show the superiority of a simple range model compared to a GARCH specification.

### 6.2.1 Theory of the range as a volatility estimator

Taking the range as an estimate for the diffusion constant of stochastic processes like the continuous random walk has a long history going back to Feller's (1951) seminal paper where he derived the asymptotic distribution of the range for the sum of independent variables using the theory of Brownian motion. The assumptions Feller used are as follows. Let  $[u_t]$  for  $t = 1, \dots, n$  be a sequence of individually identically distributed (i.i.d.) random variables with distribution  $F(u)$  with  $E(u_t) = 0$  and  $Var(u_t) = 1$ . Let then  $S_n = u_0 + \dots u_n$ ,  $M_n = \max(0, S_1, \dots S_n)$ , and  $m_n = \min(0, S_1, \dots S_n)$ . The range is then defined as  $R_n = M_n - m_n$ , which corresponds closely to our definition in Section 6.3. Feller then derives formulas for the mean and the variance of the range  $R_n$ . For details we refer to Feller (1951).

Parkinson (1980) extended the work of Feller (1951) to the case where  $Var(u_t) = D$  and  $D$  being the random walk diffusion constant. Additionally he applies the framework to the stock market and shows that the range is a far superior estimate for the diffusion constant than the traditional estimates using closing prices only. Parkinson further derives a function describing all the moments of the range distribution, which yields:

$$E(R^p) = \frac{4}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right) \left(1 - \frac{4}{2^p}\right) \zeta(p-1) (2Dn)^{p/2},$$

where  $\Gamma$  is the gamma function, and  $\zeta(x)$  is the Riemann zeta function.  $p$  has to be real and  $\geq 1$ . The first two moments are found to be:

$$E(R_n) = \sqrt{\frac{8Dn}{\pi}} \quad (6.1)$$

$$E(R_n^2) = [4\ln(2)]Dn.$$

Another contribution showing the superiority of the range estimator over the standard squared returns is the work of Martens and van Dijk (2007). They show with simulations that the so-called realized range has a lower mean-squared error than the realized variance, where the realized range is defined as the sum of all observed equidistant intra-day ranges adjusted by some factor to account for microstructure issues. For details we refer to the paper.

### 6.2.2 Extension to non i.i.d. case

One can argue that the assumption of i.i.d. increments ( $u_t$ ) is not a realistic one because there is clear evidence that the volatility of asset returns is changing over time. So, it would be nice to extend the results of Feller (1951) and Parkinson (1980) to the non i.i.d. case. Such an exercise is not trivial because the proofs used in mentioned sources rely on the i.i.d. assumption, since they proceed by reordering the observations.<sup>1</sup> In the non i.i.d. case such an arbitrary reordering is for obvious reasons not possible.<sup>2</sup> We could not find a way to solve this problem, which makes it impossible for us to come up with moments like (6.1) for the non i.i.d. case. We can only provide the reader with some intuition and heuristic reasoning to argue that the results found by Feller (1951) and Parkinson (1980) should basically remain valid (probably with some scaling parameter accounting for the dependency). Thereby, the following is by no means a formal proof but can help to get some intuition about how such a proof might proceed. An exact proof we leave for further research.

At this point we state some standard results as published in, for example, Davidson (1994) and Davidson (2000). Davidson derived the following theorem:

Let  $S_n$  be defined as above and  $u_t$  be a martingale difference sequence with  $E(u_t) = 0$  and  $E(u_t^2) = \sigma_t^2 < \infty$  with  $\bar{\sigma}_n^2 = n^{-1} \sum_{t=1}^n \sigma_t^2$ . If  $u_t$  meets the additional conditions that

$$\text{a) } n^{-1} \sum_{t=1}^n (u_t^2 - \sigma_t^2) \xrightarrow{pr} 0, \text{ and}$$

---

<sup>1</sup>For details we refer the reader to Feller (1951).

<sup>2</sup>Therefore, the proofs cannot be used to extend the results to the more general case of dependency.

b) either

- i) the sequence is strictly stationary or
- ii)  $\frac{\max_{1 \leq t \leq n} \|u_t\|_{2+\delta}}{\bar{\sigma}_n} \leq C < \infty \quad \delta > 0, \forall n \geq 1$

then  $\nu_n = \frac{\sqrt{n}\bar{u}_n}{\bar{\sigma}_n} \xrightarrow{d} \nu \sim N(0, 1)$ .

Let us also assume that

$$\frac{E(S_n^2)}{n} \rightarrow \sigma^2 < \infty \quad (\text{global wide-sense stationarity}). \quad (6.2)$$

If we then define  $X_n(r) = \frac{S[nr]}{\sqrt{n}\sigma}$ , then  $X_n \xrightarrow{d} B$ .  $\triangleright$

Here  $\xrightarrow{pr}$  and  $\xrightarrow{d}$  stand for convergence in probability and distribution, respectively.  $B$  stands for Brownian motion and  $r$  in between 0 and 1. A proof of this theorem can be found in Davidson (1994), Theorems 27.14 and 29.6.

This theorem then states that under the condition of not too strong dependence in the sequence of  $\sigma_t$  the correctly weighted partial sum  $S_n$  still converges to Brownian motion. Such a convergence is the basis of the proofs in Feller (1951) and Parkinson (1980) for the distribution of the range estimate for the diffusion constant. Heuristically speaking then, this theorem provides the basis to reason that the limit distribution of a correctly scaled range as in Equation (6.1) will stay the same with  $D = \bar{\sigma}_n^2$  even for the non i.i.d. case of the sequence  $u_t$ . Such a scaling is likely to depend on the structure of the dependency.

Another intuitive way to justify that one can use the results of Feller (1951) and Parkinson (1980) is provided in Alizadeh et al. (2002). They propose to simplify the underlying continuous time process by mapping it into a discrete time process. In practice the econometrician does not observe a continuous but a discrete time price process. One can divide the sample period  $[0, T]$  into  $n$  intervals with equal length  $L = T/n$  corresponding to say one day or one week. Then we replace the continuous volatility dynamics with a piecewise-constant process, where we assume that the volatility within every interval  $i$  is constant, i.e.  $\sigma_t = \sigma_{iL}$ . From one interval to the next we allow the volatility to change. This implies that within each interval  $i$  the price process follows a geometric Brownian motion such that Feller's (1951) and Parkinson's (1980) results apply again for that interval.

### 6.2.3 Monte Carlo experiment

As already mentioned above, Parkinson (1980) and Alizadeh et al. (2002) find that the range and log range are more efficient estimators for the diffusion constant or underlying instantaneous volatility  $D$ . These results are well-known for the i.i.d.

case. We now perform a small-scale Monte-Carlo simulation in order to show that this also holds for the non i.i.d. case. We assume a piecewise-constant volatility process but allow the volatility to change from one interval to another. The exact process we assume here is a standard GARCH(1,1) with a constant mean equation. Coefficients are based upon US stock market data and the model looks as follows:

$$\begin{aligned} r_t &= 0.0018 + \epsilon \quad \text{where} \\ \epsilon &= \sigma_t u_t \quad \text{with } u_t \sim N(0, 1). \end{aligned} \quad (6.3)$$

We assume the variance to follow a standard GARCH(1,1) model, which has become the workhorse in the volatility literature<sup>3</sup>:

$$\sigma_t^2 = 0.0000159 + 0.119\epsilon^2 + 0.851\sigma_{t-1}^2 \quad (6.4)$$

With this process we simulate  $T = 2500$  observations of  $r_t$  and  $\sigma_t^2$ , roughly corresponding to the sample length of our US stock market data. In order to be able to also simulate  $T$  range observations, we assume that during every interval the volatility is constant. We then simulate  $T_{t,int}$  intermediate returns  $r_{t,int}$  with an underlying volatility  $\sqrt{(\sigma_t^2/T_{int})}$  giving us in total  $T \times T_{int}$  returns and  $T$  variance observations. With these intra-period observations we can construct a quasi-continuous price process and obtain our  $T$  range estimates. In order to justify the usage (6.1), which is based on the assumption of Brownian motion, we take  $T_{int} = 1000$ . Larger values for  $T_{int}$  would only make the results more accurate but would not add much to the point we want to make.

In order to show the efficiency gains of the range estimator and our model we propose to use later in this chapter, we compare the fitted values of two different models. After one simulation of the outlined process we have  $T$  observations of intra-period returns and  $T$  observations of ranges. Now we would like to compare the standard GARCH model to a very basic model based upon the range.<sup>4</sup> To this end we estimate a standard GARCH(1,1) model like in (6.3) and (6.4), which also serves as the data generating process (DGP). This model is the benchmark model. In a second step we use the simulated range observations and take the log of it to arrive at the log-range. Such a monotonic transformation is done to make estimation of the model easier and will be reversed after estimation in order to arrive again at the range. We then fit a standard ARMA(1,1) model on the log-range, which has the same amount of parameters as the GARCH(1,1) model. Taking the fitted values of this model we can calculate the expected values for the

<sup>3</sup>See, for example, Ghysels et al. (2006).

<sup>4</sup>Later in this chapter we further develop such a model extending it to be Markov-switching. At this point we continue to assume that it is not Markov-switching but linear.



underlying diffusion constant of the DGP by using (6.1).

We now compare the fitted values of the GARCH(1,1) based on the noisy measure of squared returns to the fitted values of the range-based ARMA(1,1) model, by calculating the mean square errors (MSE). We just calculate for every  $t$  the difference between the model-induced estimates and the true simulated values, square them and take the average. In our exercise we do this 1000 times in order to get a distribution of MSEs for the GARCH and the for the ARMA-range model. Those two can now be compared. Results show that the MSE of the range-based model is on average 27.63% lower than that of the GARCH model. The minimum and maximum of MSE improvements were 10.79% and 45.84%, respectively.

Already this small-scale Monte-Carlo experiment shows the superiority of range-based estimates of volatility compared to squared return estimates using a standard GARCH model even though the data have been simulated by such a GARCH model. This improvement is mainly due to the fact that the range is the more efficient volatility proxy compared to the noisy proxy of squared returns. These results confirm, for example, those obtained by Parkinson (1980) and Alizadeh et al. (2002) but extend them towards the non i.i.d. case. We, therefore, propose to use the range as a basis for volatility models instead of squared inter-interval returns.

## 6.3 Methodology

Taking the results from Section 6.2 about the range as a volatility estimator into account, we now outline the general methodology proposed in this chapter. We will introduce the estimation and forecasting technique that we apply for our Markov-switching (MS) Range Model and its different specifications.

Markov-switching time series models in econometrics today draw heavily on Hamilton (1989) and Hamilton (1990) where he develops the idea that output and business cycles in an economy may be subject to discrete changes in regimes underlying their DGPs. Hamilton argues that during economic expansions the average GDP growth rate should be different compared to times of recessions and that such a behavior might best be described by a Markov chain that governs switches from regime 1, expansion say, to regime 2, recession, and vice versa. In his paper he proposes to model the GDP growth rate as a Markov-switching autoregressive process of order  $q$  (MS-AR( $q$ )).

### 6.3.1 The path to the model

Hamilton and Susmel (1994) and Cai (1994) argue that in financial time series

often observed volatility clusters or volatility persistence can be modeled in a similar fashion as in Hamilton (1989). In their paper, they develop a MS-ARCH model. ARCH models go back to the pioneering work of Engle (1982) which Bollerslev (1986) extended to generalized-ARCH (GARCH) models. Those models are designed to model the conditional second moment or variance of time series and usually fit, for example, stock market returns very well.<sup>5</sup> Hamilton and Susmel find that ARCH models often impute much persistence to stock volatility but fail to give good forecasts. They pose that this might be due to large shocks that arise from different “regimes” rather than normal shocks. One finding is that the parameters of an ARCH process seem to come from different regimes where transitions between regimes are governed by an unobserved Markov chain.

An important advantage of GARCH models over ARCH models is that they usually capture much better the time dependence in the volatility. In order to be more precise we introduce a very general  $GARCH(p, q)$  model. We refrain for the moment from specifying a mean equation but will do so in a later section. The GARCH model can be written as:

$$\begin{aligned} u_t &= \sqrt{h(\theta_h, \Phi_{t-1})} v_t \\ &= \sqrt{h_t} v_t, \end{aligned} \quad (6.5)$$

with the conditional variance of  $u_t$  specified as a function like:

$$\begin{aligned} Var(u_t) &= h_t = f(u_{t-1}, u_{t-2}, \dots) \\ &= \omega + \sum_{i=1}^p \alpha_i u_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \end{aligned} \quad (6.6)$$

where  $\theta_h$  is a vector of parameters governing the variance equation and  $v_t$  is an i.i.d. sequence with zero mean and unit variance.  $\Phi_{t-1}$  is the information set generated by  $u_t$  and represents the available information up to time  $t - 1$ . Other assumptions for the error distributions are generally possible. A  $GARCH(p, 0)$  model is equal to an  $ARCH(p)$ . So, the GARCH representation allows for a richer parametrization of the conditional variance and facilitates modeling the observed volatility persistence.

Both Cai (1994) and Hamilton and Susmel (1994) claim that the extension of GARCH processes to the Markov-switching (MS) framework is intractable because of its path-dependence. Path dependence here means that the distribution at time  $t$ , if made conditional on regime ( $S_t$ ) and on the available information set  $\Phi_{t-1}$ , depends directly on  $S_t$  but also indirectly on the whole history of regimes

<sup>5</sup>See, for example, Sabbatini and Linton (1998).

$(S_{t-1}, S_{t-2}, \dots, S_0)$ . Regime-dependence in MS-GARCH models arises through the lagged variance and lagged squared error terms. In such models, conditional variance at time  $t$  depends on the squared error and the conditional variance at time  $t - 1$ , which obviously depends on regime  $S_{t-1}$  and the squared errors and conditional variance at time  $t - 2$  and so forth. This introduces an infinite path dependence on the unobserved regimes  $S_t, S_{t-1}, \dots, S_0$  or  $\tilde{S}_t$ . In (quasi) maximum likelihood estimations (QMLE) the likelihood function could only be constructed by integrating out all possible regime paths. If we denote  $K$  as the number of regimes and  $T$  the full sample time dimension, then there are  $K^T$  possible regime-path realizations, which would make estimation impossible as the time dimension increases.

In order to avoid this path-dependence problem present in GARCH models, Gray (1996) and Klaassen (2002) introduce ways to integrate out the path-dependence inherent in GARCH models avoiding the integration over  $K^T$  possible likelihoods. Gray's idea is to integrate out the unobserved regime-path  $\tilde{S}_t$  where it emerges namely in Eq.(6.6) itself. To see this we now have to write Eq.(6.6) in a regime-dependent form:

$$\begin{aligned} \text{Var}(u_t | \tilde{S}_t, \Phi_{t-1}) &= h_{k,t} \\ &= f(u_{t-1}, u_{t-2}, \dots; \tilde{S}_t) \\ &= \omega_k + \sum_{i=1}^p \alpha_{k,i} u_{t-i}^2 + \sum_{j=1}^q \beta_{k,j} h_{k,t-j}, \end{aligned} \tag{6.7}$$

where  $\text{Var}(u_t | \tilde{S}_t; \Phi_{t-1})$  denotes the variance of  $u_t$  conditional on observable information  $\Phi_{t-1}$  and the unobservable full regime path  $\tilde{S}_t$ . The parameters in the variance equation at time  $t$  are only determined by the current regime  $S_t$ . In Eq.(6.7) there is still the full regime-path-dependence present and it is not possible to estimate its parameters.

Different ways of integrating out the path-dependence have been suggested in the literature. Hamilton and Susmel (1994) circumvent the problem of path-dependence by excluding the lagged conditional variance terms  $h_{k,t-1}, \dots, h_{k,t-q}$  in the variance equation. Hereby they only need to integrate  $K^p$  different paths out of the likelihood function in order to estimate the parameters. Gray (1996) uses a different idea. As already mentioned above he integrates out the path dependence in the GARCH by taking expectations of the conditional variances over all possible regimes. Hereby, he makes the conditional variance at time  $t$  only dependent on the current regime  $S_t$  but not the full path  $\tilde{S}_t$ . In equation form this might be

written like:

$$\begin{aligned} Var(u_t|S_t, \Phi_{t-1}) = & \omega_k + \sum_{i=1}^p \alpha_{k,i} E_{t-2} u_{t-i}^2 \\ & + \sum_{j=1}^q \beta_{k,j} E_{t-j-1} Var(u_{t-j}|S_{t-j}, \Phi_{t-j-1}), \end{aligned} \quad (6.8)$$

where  $E_{t-j-1}$  means that expectations are taken at time  $t-j-1$  over all possible regimes and conditional on the information set  $\Phi_{t-j-1}$ . This basically means that every period ex-ante probabilities are calculated (we will show the whole estimation algorithm later in this section) which are then used to weigh all possible values of  $u_{t-i}^2$  and  $Var(u_{t-j}|S_{t-j})$ . In the next period those weighted values are used as inputs for the variance equation. So, essentially the regimes  $S_{t-j}$  are integrated out at time  $t-j-1$ .

Klaassen (2002) improves on Gray's (1996) method by proposing to wait with integrating out the past regimes until they are really needed and that is at time  $t-1$ . Hereby more observations can be used in order to draw inferences about the probabilities of regimes at different points in time. If for example it is very likely that the observation at time  $t$  comes from regime  $k$  and regimes are very persistent, then this adds information to the calculation of the state probabilities in periods before. In other words Klaassen proposes to use the fact that the regime at time  $t$  essentially is in the conditioning information of  $Var(u_t|S_t, \Phi_{t-1})$  particularly if regimes are highly persistent. He, therefore, suggests to rather use the following representation:

$$\begin{aligned} Var(u_t|S_t, \Phi_{t-1}) = & \omega_k + \sum_{i=1}^p \alpha_{k,i} E_{t-1} [(u_{t-i}|S_{t-i}, \Phi_{t-i})|S_t]^2 \\ & + \sum_{j=1}^q \beta_{k,j} E_{t-1} [Var(u_{t-j}|S_{t-j}, \Phi_{t-j})|S_t], \end{aligned} \quad (6.9)$$

where the expectation  $E_{t-1}$  again is across regimes  $\tilde{S}_{t-1}$  but now conditional on the information set  $\Phi_{t-1}$  and the regime  $S_t$ . Constructed like this,  $Var(u_t|S_t, \Phi_{t-1})$  again only depends on the current regime  $S_t$  and not on the full regime path  $\tilde{S}_{t-1}$  and the path-dependence problem disappears.

Chou (2005) proposes another way to model the volatility of asset markets by using the range of the price process as an observable estimator for volatility. We already showed in Section 6.2 that the range is a very efficient volatility estimator.

Chou (2005) then suggests to model the mean of the range in the following way:

$$\begin{aligned} R_t &= \lambda_t \epsilon_t \\ \lambda_t &= \omega + \alpha R_{t-1} + \beta \lambda_{t-1}, \end{aligned}$$

where  $\epsilon_t \sim F(1, \cdot)$ . Here  $\lambda_t$  can be interpreted as the expectation of the range at time  $t$  and is modeled in an autoregressive fashion very much like a GARCH model. As can be easily seen, this model is from the multiplicative class of models and asks for an error distribution with a non-negative support in order to guarantee positivity of the range. Chou shows that this model fits the S&P500 range data quite well. Another approach is due to Alizadeh et al. (2002) who use the log-range in a stochastic volatility model.

In this paper we present a new way to combine the range volatility estimator introduced by Feller (1951) and Parkinson (1980) with the approaches of Klaassen (2002) and Chou (2005). We extend Klaassens (2002) approach to a more general MS-ARMA(a,b)-GARCH(p,q) case in order to model the log-range. As our main focus is modeling asset volatility with the help of the range estimator, we have to focus also on the mean equation and not only on the variance equation. As already mentioned above, the advantage of “observing” the volatility makes it possible to use standard time series methods, as in Chou (2005), to model it. This approach has two important benefits. First, we can essentially model the observed volatility. In the ARCH and GARCH literature, the volatility is not observed but rather derived as the conditional second moment from a series of asset returns. Second, this approach allows us to model the volatility of the volatility as a conditional second moment of the range. We do not need to estimate a conditional fourth moment as would be the case if we used return data. So, we can also model the dynamics and persistence of the volatility of the volatility of assets relatively easily.

### 6.3.2 The model

In this subsection we present the model we would like to fit to the data in its most general form. In Section 6.4 we fit different versions of such a model to the data. Let  $p_t$  denote the logarithm of the price of some speculative asset at time  $t$ . Then the range of that asset over a certain period, say a week, can be defined as  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ . Here  $p_t^{Max}$  and  $p_t^{Min}$  denote the highest and the lowest observed price of that asset over the considered time period, respectively. In other words, the range measures the maximum spread in percent of an asset’s price over a specified period. Let  $r_t$  denote the logarithm of  $R_t$ . We, thereby, use the same definition of the range and its logarithm as in Alizadeh et al. (2002). Contrary to Chou (2005) we model the log-range instead of the range in order to allow also

for negative observations. This basically changes a multiplicative into an additive model and facilitates estimation.

In the following we will use the terms range and log-range interchangeably. We only make a clear distinction at those points where it is essential. Our MS-ARMA-GARCH range model consists of four elements. First there is a mean process that governs the dynamics of the conditional mean of the range. The second element is the process for the variance specifying the dynamics of the conditional variance of the error terms. Third we have to identify the process governing the regimes. Here, we restrict ourselves to two regimes, namely a low and a high volatility regime. By allowing for two regimes only we follow contributions like Hamilton (1989), Gray (1996), Bollen et al. (2000) and Klaassen (2002). Extensions to more than two regimes are nevertheless possible. A last ingredient is the assumed error distribution. As already indicated before, the mean equation is assumed to follow an ARMA(a,b), the variance a GARCH(p,q) process and the regimes we assume to follow an unobserved first order Markov chain. We will assume the errors to be i.i.d. standard Gaussian.

Let us start with specifying the mean equation:

$$r_t = \mu_k + \sum_{i=1}^a a_{k,i} r_{t-i} + \sum_{j=1}^b b_{k,j} E_{t-1}[(\epsilon_{t-j} | S_{t-j}, \Phi_{t-j}) | S_t] + \epsilon_t, \quad (6.10)$$

where

$$\begin{aligned} \epsilon_t &= \sqrt{h_{k,t}} z_t \\ h_{k,t} &= \omega_k + \sum_{m=1}^p \alpha_{k,m} E_{t-1}[(\epsilon_{t-m} | S_{t-m}, \Phi_{t-1}) | S_t] \\ &\quad + \sum_{n=1}^q \beta_{k,n} E_{t-1}[(h_{t-n} | S_{t-n}, \Phi_{t-1}) | S_t]. \end{aligned} \quad (6.11)$$

In the mean equation  $\mu_k$  represents the constant term for all different regimes  $k = 1, 2, \dots, K$ ,  $a_{k,i}$  are all autoregressive coefficients,  $b_{k,j}$  are all moving average coefficients and  $z_t$  is assumed to be i.i.d. with a  $N(0, 1)$  distribution. In Eq.(6.11)  $\omega_k$  is the constant term of the variance equation,  $\alpha_{k,m}$  and  $\beta_{k,n}$  are the lagged squared error and lagged variance coefficients, respectively. By this it is clear that  $S_t$  fully determines the parameters of the conditional distribution of  $r_t$ .

As, for example, in Hamilton (1989) we assume that the regimes  $S_t$  follow a

first-order Markov process with constant transition probabilities<sup>6</sup>

$$p(S_t = j | S_{t-1} = i, S_{t-2} = j, \dots; \Phi_{t-1}) = p(S_t = j | S_{t-1} = i) = p_{ij}, \quad (6.12)$$

for  $i, j = 1, 2, \dots, K$ . So, as required by the Markov property the probability of state  $S_t = j$  only depends on  $S_{t-1}$ , namely the state the process was in at time  $t-1$ . All these probabilities can be summarized in a  $(K \times K)$  transition probability matrix:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & \dots & p_{K1} \\ p_{12} & p_{22} & \dots & p_{K2} \\ \vdots & \vdots & \dots & \vdots \\ p_{1K} & p_{2K} & \dots & p_{KK} \end{bmatrix},$$

where each column of  $\mathbf{P}$  sums to unity.

### 6.3.3 Estimation

In the regime-switching literature, models are usually estimated by quasi maximum likelihood (QMLE). Gray (1995) proves for some regime-switching models the consistency and asymptotic normality of the QML estimator under relatively mild regularity conditions. We, therefore, follow this path with our MS-ARMA-GARCH range model. As in Gray (1996) and Klaassen (2002), our likelihood has a first-order recursive structure and can be estimated similar to a normal single regime GARCH model. At the same time one can calculate probabilities that the process is in a particular regime at a specific time  $t$ , which is very useful if we want to classify our series into periods with low and high volatility. Also following Gray and Klaassen we use two different types of regime probabilities. The first is the ex ante probability of a certain regime. It will be denoted as  $p(S_t | \Phi_{t-1})$  and is the conditional probability that the process is in a certain regime at time  $t$  given only the information set available at time  $t-1$ . Second, we also calculate the smoothed regime probabilities  $p(S_t | \Phi_T, \theta)$  or in short  $p(S_t | \Phi_T)$  which use the complete data and information set  $\Phi_T$  at the estimated coefficient vector  $\theta$ , thereby smoothing the ex ante probabilities. These smoothed regime probabilities give the econometrician's best inference about the probability of the regime the process was in at time  $t$  and will be calculated from the ex ante probabilities we obtain during estimation of the model.

We now introduce the estimation procedure by extending the work of Klaassen (2002) and Gray (1996) to the general case of a MS(K)-ARMA(a,b)-GARCH(p,q)

---

<sup>6</sup>In general it is also possible to model the transition probabilities as time-varying. Examples are the contributions of Diebold, Lee, and Weinbach (1994) and Gray (1996).

model. Klaassen and Gray are mostly concerned with the Markov-switching aspects in the conditional variance equation. We observe an estimator of the variance and are rather focussing on the mean equation of course not neglecting the variance of the process. Above in Eq.(6.10) and (6.11) we already presented a more general model essentially using the same ideas as in Klaassen's paper. Now we turn to the estimation procedure for those models.

In order to obtain the full sample likelihood function we basically have to model the density of every range observation at time  $t$  for all possible regimes conditional on only observable information. So, we write that density as:

$$\begin{aligned} f(r_t|\Phi_{t-1}) &= \sum_{k=1}^K f(r_t, S_t = k|\Phi_{t-1}) \\ &= \sum_{k=1}^K f(r_t|S_t = k, \Phi_{t-1})p(S_t = k|\Phi_{t-1}), \end{aligned} \quad (6.13)$$

where we take the sum  $\sum_{k=1}^K$  of the regime conditional densities over all possible regimes weighted by their respective ex ante probabilities of occurrence  $p(S_t = k|\Phi_{t-1})$ . Therefore, we can write the distribution of  $r_t$  conditional on available information like:

$$r_t|\Phi_{t-1} \sim \begin{cases} f(r_t, S_t = 1|\Phi_{t-1}) \text{ with probability } p(S_t = 1|\Phi_{t-1}), \\ f(r_t, S_t = 2|\Phi_{t-1}) \text{ with probability } p(S_t = 2|\Phi_{t-1}), \\ \vdots \\ f(r_t, S_t = K|\Phi_{t-1}) \text{ with probability } p(S_t = K|\Phi_{t-1}). \end{cases}$$

In the empirical section of this paper we restrict ourselves to the case of  $K = 2$ . If we assume conditional normality for the error distribution in Eq.(6.10) we can write:

$$f(r_t|S_t = k, \Phi_{t-1}) = \frac{1}{\sqrt{2\pi h_{k,t}}} \exp \left\{ \frac{-\epsilon_{k,t}^2}{2h_{k,t}} \right\}. \quad (6.14)$$

In general, the errors can for example also be assumed to follow a student-t distribution obviously changing (6.14) accordingly.

As in Gray (1996) and Klaassen (2002) and according to the assumed first-order Markov structure, the probability  $p(S_t = k|\Phi_{t-1})$  depends only on the regime the whole process is in at time  $t - 1$ . If we condition on the regime at time  $t - 1$  one



can write the ex-ante probability as:

$$\begin{cases} p(S_t = 1|\Phi_{t-1}) = \sum_{k=1}^K p(S_t = 1|S_{t-1} = k, \Phi_{t-1})p(S_{t-1} = k|\Phi_{t-1}), \\ p(S_t = 2|\Phi_{t-1}) = \sum_{k=1}^K p(S_t = 2|S_{t-1} = k, \Phi_{t-1})p(S_{t-1} = k|\Phi_{t-1}), \\ \vdots \\ p(S_t = K|\Phi_{t-1}) = \sum_{k=1}^K p(S_t = K|S_{t-1} = k, \Phi_{t-1})p(S_{t-1} = k|\Phi_{t-1}), \end{cases} \quad (6.15)$$

where, according to the Markov property,

$$p(S_t = j|S_{t-1} = i, \Phi_{t-1}) = p(S_t = j|S_{t-1} = i) = p_{ij}. \quad (6.16)$$

So, the probabilities  $p(S_t = j|S_{t-1} = i, \Phi_{t-1})$  only depend on  $S_{t-1}$  and are equal to the fixed transition probabilities in Eq.(6.12) which are summarized in the transition matrix  $\mathbf{P}$ .

Further note that the second part on the right hand side of Eq.(6.15),  $p(S_{t-1} = k|\Phi_{t-1})$  we can write, according to Bayes' Rule, as:

$$\begin{aligned} p(S_{t-1} = k|\Phi_{t-1}) &= p(S_{t-1}|r_{t-1}, \Phi_{t-2}) \\ &= \frac{p(r_{t-1}|S_{t-1}, \Phi_{t-2})p(S_{t-1}|\Phi_{t-2})}{p(r_{t-1}|\Phi_{t-2})} \\ &= \frac{p(r_{t-1}|S_{t-1}, \Phi_{t-2}) \sum_{S_{t-2}=1}^K p(S_{t-2}|\Phi_{t-2})p(S_{t-1}|S_{t-2})}{p(r_{t-1}|\Phi_{t-2})}. \end{aligned} \quad (6.17)$$

Here, the variables needed to compute  $p(S_{t-1} = k|\Phi_{t-1})$  are its previous values  $p(S_{t-2} = k|\Phi_{t-2})$ , the constant transition probabilities  $p_{ij}$  and the densities  $p(r_{t-1}|S_{t-1}, \Phi_{t-2})$  and  $p(r_{t-1}|\Phi_{t-2})$  from the same calculation one step before. So, the computation of  $p(S_{t-1} = k|\Phi_{t-1})$  is a first-order recursive process.

Before we move to the last ingredient for the calculation of the likelihood function we present the full-sample log-likelihood which can be obtained as:

$$\mathbf{L} = \sum_{t=\max(a,b,p,q)}^T f(r_t|S_t)p(S_t|\Phi_{t-1}) \quad (6.18)$$

where the first part on the right hand side is the conditional density at time  $t$  given in Eq.(6.14) and where the second part is the ex ante regime probability described in Eq.(6.15). Unfortunately, the density  $f(r_t|S_t)$  cannot be calculated in a straightforward fashion because of the path dependency in the moving average and variance part of a plain ARMA-GARCH model. So, we have to use Eq.(6.10) and (6.11) which necessitates the calculation of the expectations of lagged error

and variance terms across regimes. Klaassen (2002) proposed to use all available information up to time  $t - 1$  to calculate the expected lagged variance in the variance equation of a Markov-switching GARCH(1,1) model. We propose to use the same probability measure to also weigh the lagged error terms in the MA-part and the lagged squared errors in the ARCH-part of our proposed model. In his paper Klaassen proposes a weighing mechanism which gives the probability that the previous regime was  $S_{t-1}$  given that the current regime is  $S_t$  and given the information set  $\Phi_{t-1}$ . It can be stated in the following way:

$$p(S_{t-1}|S_t, \Phi_{t-1}) = \frac{p(S_{t-1}|\Phi_{t-1})p(S_t|S_{t-1}, \Phi_{t-1})}{p(S_t|\Phi_{t-1})} \quad (6.19)$$

$$= \frac{p(S_{t-1}|\Phi_{t-1})p_{ij}}{p(S_t|\Phi_{t-1})}, \quad (6.20)$$

where  $p(S_{t-1}|\Phi_{t-1})$  is given by Eq.(6.17),  $p_{ij}$  are the fixed transition probabilities in Eq.(6.12), and  $p(S_t|\Phi_{t-1})$  is given by Eq.(6.15). If one wishes to estimate models with a lag structure  $max = max(b, p, q) > 1$  one obviously needs the corresponding probabilities  $(p(S_{t-2}|S_t, \Phi_{t-1}), p(S_{t-3}|S_t, \Phi_{t-1}), \dots, p(S_{t-max}|S_t, \Phi_{t-1}))$ , in order to get the expected values of those lagged error and variance terms as well. Those probabilities can be calculated in a similar way as in Eq.(6.19). This completes the description of the estimation procedure.

### 6.3.4 Smoothed regime inference

As mentioned above, the smoothed regime probabilities represent the econometricians best inference about the regime the process was in at time  $t$  using all available information up to time  $T$ .<sup>7</sup> In general, one can write  $p(S_t|\Phi_\tau)$  for all  $K$  regimes the ex post probability as:

$$\begin{aligned} p(S_t|\Phi_\tau) &= p(S_t|r_\tau, \Phi_{\tau-1}) \\ &= \frac{p(r_\tau|S_t, \Phi_{\tau-1})p(S_t|\Phi_{\tau-1})}{\sum_{S_t=1}^K p(r_{\tau-1}|S_t, \Phi_{\tau-1})p(S_t|\Phi_{\tau-1})}. \end{aligned} \quad (6.21)$$

When  $\tau = t$ , then  $p(S_t|\Phi_\tau)$  follows directly because we already know  $p(S_t|\Phi_{\tau-1})$  and  $p(r_{\tau-1}|S_t, \Phi_{\tau-1})$  from the foregoing maximum likelihood estimation process. For all the following times ( $\tau = t + 1, t + 2, \dots, T$ ), the calculation of the smoothed probabilities is a first-order recursive process.

If  $\tau > t$  we basically need two inputs in order to compute Eq.(6.21). The first in-

<sup>7</sup>This section heavily draws on results by Hamilton (1989), Hamilton (1990), Gray (1996) and especially Klaassen (2002).

gredient is the previous  $K$  ex post probabilities  $p(S_t|\Phi_{\tau-1})$ , which are known from the previous iteration. Second, we need to compute the density  $p(r_{\tau-1}|S_t, \Phi_{\tau-1})$  for all  $K$  possible regime outcomes. In order to arrive at this density we have to go through some steps. First we can write it as:

$$p(S_\tau|r_t, \Phi_{\tau-1}) = \sum_{S_\tau=1}^K p(r_\tau|S_\tau, \Phi_{\tau-1})p(S_\tau|S_t, \Phi_{\tau-1}), \quad (6.22)$$

where one uses the insight that the conditional distribution of  $r_\tau$  given  $S_\tau$  does not depend on the earlier regimes  $(S_t, S_{t-1}, \dots)$  because we integrate out the path-dependence during the estimation procedure. In Eq.(6.22) we again have two parts on the right hand side. The first one contains the densities  $p(r_\tau|S_\tau, \Phi_{\tau-1})$  for all  $K$  regimes, which are known from the estimation procedure. The second part,  $p(S_\tau|S_t, \Phi_{\tau-1})$ , consists of the  $\tau - t$  period transition probabilities of the Markov chain for all possible regime outcomes. By using the Markov property we can rewrite it as:

$$\begin{aligned} p(S_\tau|S_t, \Phi_{\tau-1}) &= \sum_{S_{\tau-1}=1}^K p(S_\tau|S_{\tau-1}, \Phi_{\tau-1})p(S_{\tau-1}|S_t, \Phi_{\tau-1}) \\ &= \sum_{S_{\tau-1}=1}^K p_{ij}p(S_{\tau-1}|S_t), \end{aligned} \quad (6.23)$$

where we again, as a first ingredient, have the one period ahead  $p_{ij}$  transition probabilities following from Eq.(6.12). The second part on the right hand side on Eq.(6.23) can be calculated recursively.

Let us write  $p(S_{\tau-1}|S_t, \Phi_{\tau-1})$  for all  $K^2$  regime combinations as:

$$\begin{aligned} p(S_{\tau-1}|S_t, \Phi_{\tau-1}) &= p(S_{\tau-1}|S_t, r_{\tau-1}, \Phi_{\tau-2}) \\ &= \frac{p(r_{\tau-1}|S_{\tau-1}, \Phi_{\tau-2})p(S_{\tau-1}|S_t, \Phi_{\tau-2})}{\sum_{S_{\tau-1}=1}^K p(r_{\tau-1}|S_{\tau-1}, \Phi_{\tau-2})p(S_{\tau-1}|S_t, \Phi_{\tau-2})}, \end{aligned} \quad (6.24)$$

where we use the fact that the conditional density  $p(r_{\tau-1}|S_{\tau-1}, \Phi_{\tau-2})$  is independent of all earlier regimes once  $S_{\tau-1}$  is given. For iteration  $\tau$  all ingredients in Eq.(6.24) are known either from the foregoing estimation procedure (the conditional density  $p(r_{\tau-1}|S_{\tau-1}, \Phi_{\tau-2})$ ) or the previous iteration in the calculation of the smoother (the  $(\tau-t-1)$ -period ahead transition probability  $p(S_{\tau-1}|S_t, \Phi_{\tau-2})$ ). The ex post probability for  $\tau = T$  then gives the smoothed regime probability  $p(S_t|\Phi_T)$ , which completes the calculation of the smoothed probabilities.

## 6.4 Application and results

In this section of the chapter, we present the results of fitting our model in Eq.(6.10) and (6.11) to the data. First we concentrate ourselves on the S&P500. Later we extend the analysis to the ten year T-notes, three months T-bills, the FTSE100, and the Nikkei225. As a first step we present the data themselves, some descriptive statistics and evidence indicating that there very well might be a hidden Markov process underlying the data causing the data generating process to switch between a low and a high volatility state. As already mentioned above, we assume a two regime Markov process. We then further present the results of fitting different versions of our MS-ARMA-GARCH range model and find that model which fits the data best. We will end this section by briefing possible interpretations of the results.

### 6.4.1 The data

The data we use are weekly ranges for the US stock market index S&P500 from January 2<sup>nd</sup> 1962 until February 11<sup>th</sup> 2008 summing to 2406 observations in total (observations are on Mondays). We downloaded them from the *yahoo.com* database. In order to arrive at the actual data we transformed the downloaded  $p_t^{Max}$  and  $p_t^{Min}$ , the highest and the lowest (log-)price index observation, respectively, like:

$$R_t = 100 * (p_t^{Max} - p_t^{Min}). \quad (6.25)$$

The range  $R_t$  is by definition a positive variable and would ask for either a multiplicative model and/or an error distribution that has a lower bound at zero. Furthermore, its unconditional distribution is highly skewed further complicating its modeling. We, therefore, use the log-range<sup>8</sup>:

$$r_t = \ln(R_t), \quad (6.26)$$

which unconditional distribution is surprisingly close to a normal distribution. This result confirms results of, for example, Alizadeh et al. (2002) who also find that the log-range can very well be described as normally distributed - a fact that is uncommon in financial time series, which are usually skewed and show

<sup>8</sup>In fact we use an outlier-adjusted version of the data series. We consider all realizations as outliers that are in the upper and the lower 1% percentile of the unconditional distribution. Less than five trading days per week are often responsible for lower tail outliers. Identified outliers are eliminated by taking the average of five consecutive observations, namely the two observations before and after the outlier and the outlier observations itself. By this method we make sure that extreme observations remain extreme. A robustness check showed no significant changes (besides larger Jarque Bera test statistics) in estimation results.

excess kurtosis. Furthermore, Andersen et al. (2003) find that forecasting the log transformation of volatility yields better in- and out-of sample forecasts of the variance because it puts less weight on extreme realizations of the volatility.

Table 6.1 shows descriptive statistics of our range and log-range data. The Jarque-Bera test statistics reject normality for both series. In the case of log-ranges the statistic still rejects normality but is already very much closer to non-rejection than in the case of the range. We also perform an augmented Dickey and Fuller (1979) test (ADF) with lag-length selection using the Schwarz (1978) information criterion. The null hypothesis of a unit root is clearly rejected for all series. So, there is no need for taking the first difference of the data.

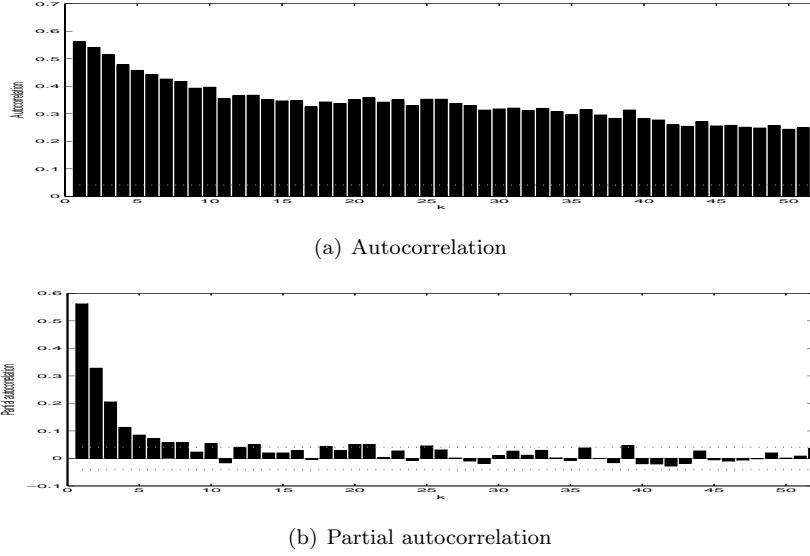
Table 6.1: Descriptive statistics

	Weekly observations	
	Range	Log-Range
Mean	3.145	1.045
Median	2.745	1.010
Maximum	12.215	2.503
Minimum	0.988	-0.012
Std.Dev.	1.530	0.441
Skewness	1.576	0.295
Kurtosis	6.250	2.825
Jarque-Bera	2015.069	37.173
P-value	0.000	0.000
ADF test	-9.203	-8.574
P-value	0.000	0.000

**Note:** Descriptive statistics relating weekly range and log-range observations as derived from Eq.(6.25) and (6.26) respectively. Data are from January 2<sup>nd</sup> 1962 until February 11<sup>th</sup> 2008 summing to 2406 observations in total. The data are plotted in Fig. 6.2. Augmented Dickey-Fuller (ADF) test statistics and p-values are calculated based on an automatic lag-length selection using the Schwarz information criterion.

In Figure 6.1 we show the autocorrelation function and the partial autocorrelation function for the lags  $k = 1, \dots, 52$ , respectively. We observe a quite slow decay in the autocorrelation of the S&P500 range data, which might be interpreted as long memory or, in other words, as a fractionally integrated data generating process. Nevertheless, there is the possibility of confusing long memory with structural breaks or regime-switching behavior. For contributions linking structural breaks and fractional integration see for example Bhattacharya et al. (1983), Künsch (1986), and Teverovsky and Taqqu (1997). Even more relevant here is the work of Diebold and Inoue (2001). They find theoretically and by means of simulation that “...structural change in general, and stochastic regime switching in particular, are intimately related to long memory and easily confused with it, so long

Figure 6.1: S&amp;P500 (partial) autocorrelation



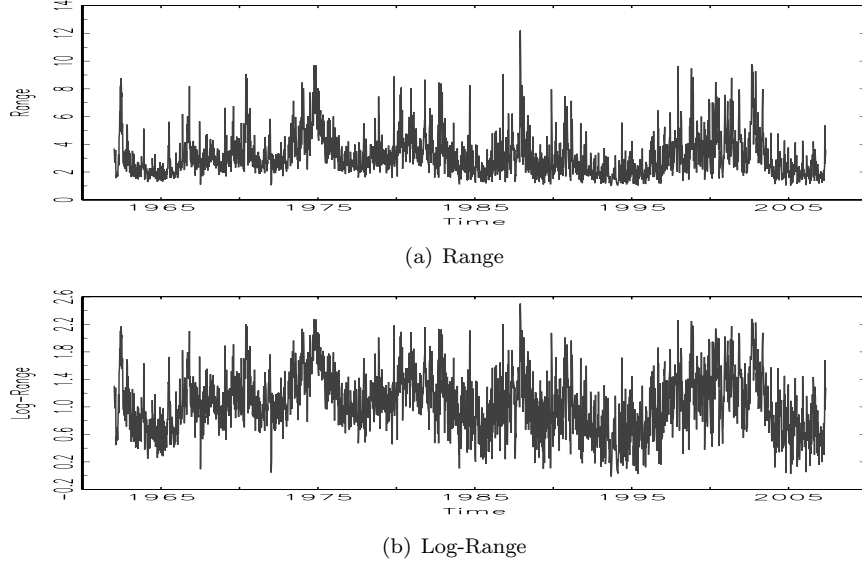
**Note:** Dotted lines represent the two standard deviation critical values.

as only a small amount of regime switching occurs in an observed sample path” (Diebold and Inoue (2001, p.153)). Anticipating on our results, we find such a “small amount” of regime switches in the data. Furthermore, we are interested in the co-occurrence of high and low volatility periods across international financial markets. We also believe that investors are likely to be driven by psychological factors more in line with regime-switching behavior than by long-memory processes. Therefore, we aim for a means of volatility regime classification and thus opt for a Markov-switching model representation as outlined before.

We show the range  $R_t$  and the log-range  $r_t$  time series in Figure 6.2. The unconditional distributions of the range and log-range are shown in Figures 6.3. Obviously, the observation from Table 6.1 that  $r_t$  is close to normally distributed makes  $R_t$  appear to have a log-normal distribution.

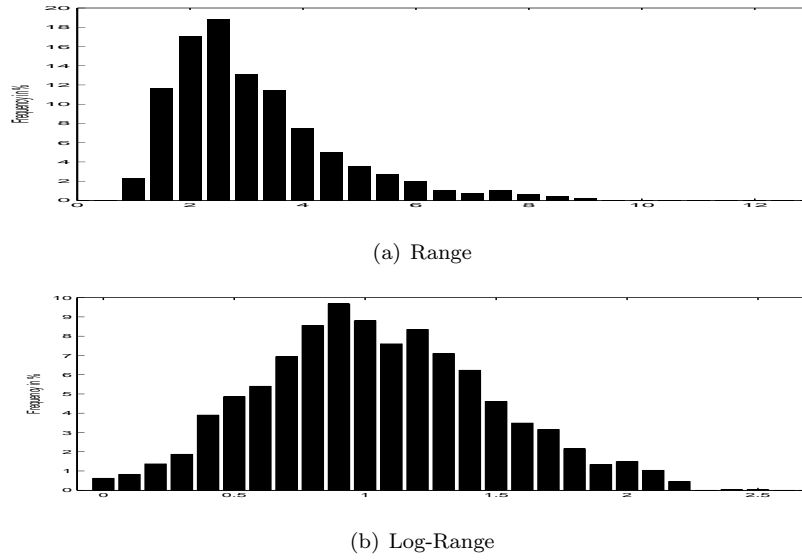
We continue the data description with an informal time series analysis by having a closer look at the data. One can see in Figure 6.2 quite clearly that there are periods of relatively low volatility and periods of high volatility. Especially the periods in the middle of the 1970s, the beginning of the 1980s, the late 1980s and from 1998 until 2003 are marked by clearly higher average volatilities measured by the range and/or log-range. High volatility in the early and middle of the 1970s coincides with the break down of the Bretton Woods gold system and the first oil

Figure 6.2: S&P500 range and log-range



**Note:** Range is equal to  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ . The Log-Range is calculated as  $r_t = \ln(R_t)$ .

Figure 6.3: S&P500 range and log-range unconditional distributions



**Note:** The range and log-range are calculated as in Figure 6.2.

crisis starting in 1973, which was followed by strong reactions of world financial markets. The high volatility period in the beginning of the 1980s corresponds to the second oil crisis, where between 1980 and 1981 the price of crude oil more than doubled within a period of 12 months. In the late 1980s there is another very pronounced but relatively short period of high volatility with a pronounced peak corresponding to “Black Monday” on October 19<sup>th</sup> 1987. On this day the main US stock markets dropped by ca. 23%, starting a period with extreme uncertainty in asset markets worldwide.<sup>9</sup> This period of increased asset market volatility did not last very long though and markets returned to pre-crash volatility levels before showing some increased volatility again in the beginning of the 1990s during and after the second Gulf War. A further period of higher than normal volatility starts in 1998/1999 probably corresponding to the Russian crisis and the build-up and later burst of the “dot-com-bubble”. This period lasts until 2003 which roughly corresponds to the end of the third Gulf War.

In summary, there appear to be quite distinct periods of high and low market uncertainty corresponding to high and low volatility, as measured by the range and log-range, respectively. We think that this is strong evidence for an underlying regime-switching process that might very well be described as a Markov chain. In order to formally test for the presence of a low and high volatility regime we use the testing procedure introduced by Cheung and Erlandsson (2005). They propose a Monte Carlo based testing procedure to simulate an empirical finite sample test statistic for the null hypothesis of one regime (no Markov-switching) against the alternative of two regimes. Such a testing procedure comes in handy because standard statistical procedures fail here. Under the null hypothesis of a linear model with only one regime the nuisance parameters P11 and P22, which are present under the alternative, are not defined making the distribution of the asymptotic log-likelihood ratio test statistic non-standard. Contributions like Hansen (1992; 1996) and Garcia (1998) derived such asymptotic distributions. But still not much is known about their finite sample behavior. We therefore opt for the procedure proposed by Cheung and Erlandsson (2005) which they show to have good power also in finite samples.

We apply the Cheung and Erlandsson (2005) testing procedure to the data by comparing the best fitting linear model with the best fit of the Markov-switching models only allowing for a change in the intercept of the mean equation.<sup>10</sup> We indeed find significant results for the presence of at least two regimes. The p-

<sup>9</sup>See, for example, Shiller (1989) and Carlson (2007).

<sup>10</sup>Such a test can easily be applied to different alternative model specifications. Nevertheless, it appears sufficient to us at this point to take the simplest Markov-switching as an alternative model because it already showed up to be sufficient to generate significant results. Furthermore, any more complicated alternative model specification would have increased computing time without giving more insights.



value of the likelihood ratio test was found to be at 2%, clearly rejecting the null hypothesis of a linear specification in favor of the alternative Markov-switching hypothesis justifying the further procedure of modeling the log-range according to Eq.(6.10) and (6.11).

Another criterion for a well fitting regime-switching model should be that it is capable of at least also identifying some of the periods of high and low volatility visually found in the graphs before. Therefore, we present the results of fitting the considered MS-ARMA-GARCH models to the data in Section 6.4.2.

### 6.4.2 Estimation results

In Section 6.4.1 we showed that the weekly S&P500 range and log-range are very likely to be drawn from at least two different densities and thereby from more than one volatility regime. In this section we aim at finding the best fitting, parsimonious model from our proposed class of MS(2)-ARMA(a,b)-GARCH(p,q) models generally described in Eq.(6.10) and (6.11), which are reproduced here for convenience:

$$r_t = \mu_k + \sum_{i=1}^a a_{k,i} r_{t-i} + \sum_{j=1}^b b_{k,j} E_{t-1}[(\epsilon_{t-j} | S_{t-j}, \Phi_{t-j}) | S_t] + \epsilon_t,$$

where

$$\begin{aligned} \epsilon_t &= \sqrt{h_{k,t}} z_t \\ h_{k,t} &= \omega_k + \sum_{m=1}^p \alpha_{k,m} E_{t-1}[(\epsilon_{t-m} | S_{t-m}, \Phi_{t-1}) | S_t] \\ &\quad + \sum_{n=1}^q \beta_{k,n} E_{t-1}[(h_{t-n} | S_{t-n}, \Phi_{t-1}) | S_t]. \end{aligned}$$

We will have to fit different specifications of the MS(2)-ARMA(a,b)-GARCH(p,q) models in order to be able to decide upon which one fits the data best. We will proceed in a bottom-up way. We start with MS(2)-ARMA(a,b) specifications without looking at possible GARCH or volatility of volatility clustering effects. In order to make sure that the QMLE estimation arrived at the global maximum likelihood, we estimate the models with 100 different randomly drawn starting values. To check for a good fit we will employ different means. A very important criterium will obviously be to check, if there is any autocorrelation in the standardized residuals and/or the squared standardized residuals left. Any remaining autocorrelation in the residuals asks for an increase in the amount of ARMA-terms.

Any remaining autocorrelation in the square of the standardized residuals hints at GARCH-effects not sufficiently accounted for by the model, and we might need to add more ARCH or GARCH terms. The best fitting model will not have any remaining autocorrelation in the standardized residuals or squared standardized residuals. So, the following subsections analyze the data in more detail.

#### **Only the intercept changes with the regime**

In the empirical implementation we allow different parts of Eq.(6.10) and (6.11) to change with regimes. An ARMA-C or an ARMA-X specification mean that only the constant or all parameters in that part of the model are allowed to change, respectively. In this subsection we concentrate on the different MS(2)-ARMA-C(a,b)-GARCH(p,q) model specifications. In all the coming models we let only the constant or intercept,  $\mu_k$ , in the mean equation (Eq.(6.10)) change with the regime. Later, we also experiment with regime dependent ARMA and GARCH parameters in order to find out if the volatility of volatility is changing with time as well. We present all relevant estimation results in Table 6.2. The columns represent all different specifications with parameter estimates and standard errors reported. We also show the value of the maximized log-likelihood function and Ljung-Box (LB) and Jarque-Bera statistics in order to check for residual and squared standardized residual autocorrelation and normality of the residual distributions.

We start with the most parsimonious specification being the MS(2)-AR(1) model. Here, we can already see that there are clearly two different volatility regimes in the S&P500 data over the considered sample period. The constant terms in either regime (regime 1 and 2) differ significantly from each other. Checking for correct model specification by inspecting the Ljung-Box statistics both for the standardized residuals and squared residuals it becomes apparent that the simple MS(2)-AR(1) specification cannot completely eliminate autocorrelation in the residuals and their squares. Two points arise from this. First, we need to increase the order of ARMA-terms in the mean equation. Second, there is evidence for conditional heteroscedasticity in the residuals asking for the inclusion of some ARCH and/or GARCH terms in order to allow for a time-varying variance. Though, before specifying the conditional variance of the range, we first proceed in finding an ARMA-specification that is able to account for the autocorrelation in the residuals. Afterwards we continue with modeling the conditional heteroscedasticity.

Table 6.2: Estimation results for weekly S&P500 data only allowing the constant ( $\mu$ ) to change

Parameters	AM(1,0)		AM(1,1)		AM(1,1)-G(1,0)		AM(1,1)-G(1,1)		AM(1,1)-G(2,1)	
	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.
$\mu_1$	0.5413	0.0004	0.0410	0.0073	0.0589	0.0050	0.0560	0.0053	0.0545	0.0020
$\mu_2$	0.8865	0.0009	0.1741	0.0554	0.0934	0.0114	0.0900	0.0148	0.0876	0.0041
$a_1$	0.3426	0.0003	0.9522	0.0030	0.9236	0.0096	0.9263	0.0123	0.9280	0.0032
$b_1$			-0.7533	0.0046	-0.6956	0.0008	-0.6808	0.0094	-0.6907	0.0039
$\omega$	0.1129	0.0000	0.1051	0.0007	0.1000	0.0318	0.0003	0.0000	0.0001	0.0001
$\alpha_1$					0.0939	0.0971	0.0128	0.0005	0.0892	0.0111
$\alpha_2$							0.9842	0.0007	-0.0797	0.0109
$\beta_1$	0.9899	0.0000	0.9779	0.0196	0.9958	0.0003	0.9964	0.0011	0.9893	0.0002
$P_{11}$	0.9867	0.0004	0.6789	0.2956	0.9969	0.0000	0.9947	0.0002	0.9961	0.0003
$P_{22}$									0.9962	0.0002
Log-Likelihood	-838.060		-728.313		-717.875		-695.227		-688.276	
P-Values										
$LB_1$	0.000		0.513		0.975		0.625		0.781	
$LB_5$	0.000		0.286		0.726		0.775		0.808	
$LB_{10}$	0.000		0.524		0.871		0.909		0.917	
$LB_1^2$	0.000		0.000		0.612		0.000		0.969	
$LB_5^2$	0.000		0.000		0.003		0.002		0.375	
$LB_{10}^2$	0.000		0.000		0.000		0.007		0.318	
Jarque-Bera	13.314		6.774		36.263		33.699		32.855	
P-value	0.001		0.034		0.000		0.000		0.000	

**Note:** AM(a,b)-G(p,q) is short for an ARMA(a,b)-GARCH(p,q) specification. Parameters are estimated using the GAUSS6.0 conditional optimization package (co) under the constraints of all ARMA and GARCH roots lying outside the unit circle. Additionally, we impose a positivity constraint for the variance and conditional variance. We apply the standard convergence criteria. The parameters are as in Eq.(6.10) and (6.11) for the respective model specifications.  $P_{11}$  and  $P_{22}$  are the Markov chain transition probabilities for every period for staying in the low and in the high volatility regime, respectively.  $LB_x$  stands for the Ljung-Box test at  $x$  lags of the standardized residuals.  $LB_x^2$  is the same but for squared standardized residuals. For the Ljung-Box test we only report p-values for the null hypothesis of no autocorrelation. Ljung-Box test statistics are not adjusted for ARCH effects as suggested in Diebold (1986). The Jarque-Bera test tests for standard normality in the standardized residuals. For the Jarque-Bera test we report the test statistics and corresponding p-values.

Already an ARMA(1,1) specification is able to deliver insignificant autocorrelation levels in the residuals, which can be checked by looking at the Ljung-Box statistics.<sup>11</sup> But we still have clear evidence for remaining conditional heteroscedasticity. The addition of an equation specifying the conditional variance solves this problem.

In order to take care of the conditional heteroscedasticity in the data we specify different GARCH(p,q) models. We only report the results for the GARCH(1,0), GARCH(1,1) and the GARCH(2,1) cases in Table 6.2. The GARCH(1,0) specification for the variance equation does not seem to be sufficient to justify the i.i.d. assumption for the residuals because the Ljung Box statistics for the standardized squared residuals are still significant.<sup>12</sup> We therefore try two different approaches, namely augmenting the conditional variance with a lagged conditional variance term (GARCH(1,1)) and augmenting it with higher order ARCH-terms (GARCH(2,1)). Also the assumption of a GARCH(1,1) specification does not fully solve the problem of not having i.i.d. residuals because the Q-statistic at one lag is still significant at a 5% level. The GARCH(2,1) model though delivers insignificant autocorrelations for the squared residuals at a 10% significance level.<sup>13</sup> We therefore consider the MS(2)-ARMA-C(1,1)-GARCH(2,1) model as fitting the data best. By inspecting the Jarque-Bera test statistic it is apparent that the normality assumption is likely to be violated, though.

By having a closer look at the coefficients of the MS-ARMA-C(1,1)-GARCH(2,1) model in Table 6.2, one can see a quite clear difference in the constant terms of either regime. In the low volatility regime  $\mu_1$  is equal to 0.0545 whereas in the high volatility regime  $\mu_2$  is equal to 0.0876. These intercepts and the AR-coefficient of 0.928 give us the unconditional log-range values of 0.757 and 1.217 for the low and the high volatility regime, respectively. Such log-range values translate into ranges of 2.132 and 3.376, respectively, which corresponds to a, on average, 61% larger volatility during periods with high volatility as compared to those periods with low volatility. Furthermore, the parameters in the variance equation  $\alpha_1$  and  $\beta_1$  add up to 0.9988, which suggest a quite persistent conditional volatility of the log-range,

<sup>11</sup>The Ljung-Box statistics in Table 6.2 and Table 6.3 have been calculated without accounting for ARCH effects. Diebold (1986) suggested to adjust the standard Ljung-Box statistics in the presence of ARCH effects. Such an adjustment would nevertheless only decrease the test statistics and increase the p-values to some extent. Such an adjustment is therefore not likely to affect the choice of the appropriate model significantly and we therefore only report the unadjusted Ljung-Box test statistics. Adjusted test statistics as suggested in Diebold (1986) are available upon request.

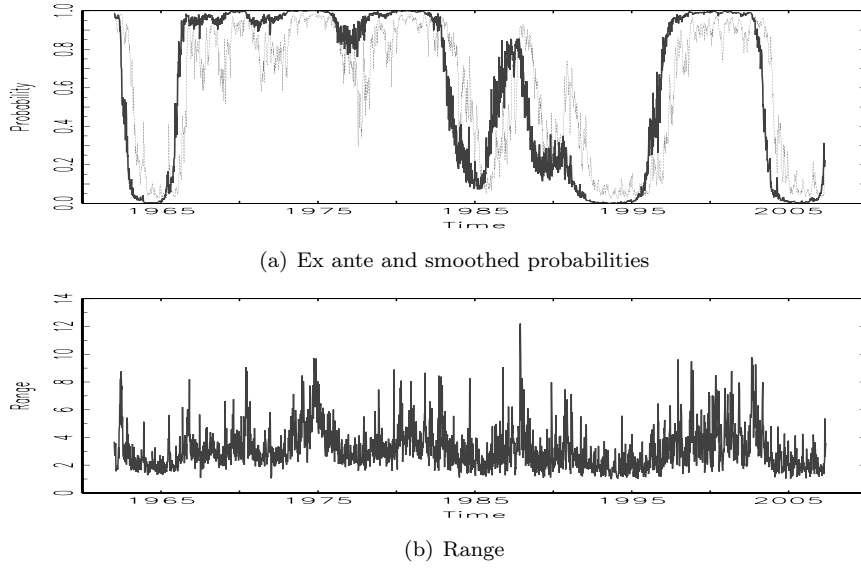
<sup>12</sup>By inspecting the Ljung-Box Q-statistics more closely, we find that especially the first four lags cause the rejection of the no autocorrelation null hypothesis. Detailed results are not reported here, but are available upon request.

<sup>13</sup>Because of space considerations we do not report all those test statistics in Table 6.2. They are available on request.

where the biggest contribution of this persistence comes from the GARCH and not the instantaneous ARCH parameters. This suggests that shocks to the volatility of the volatility die out quite slowly.

We also present the ex ante and smoothed regime probabilities derived from the best fitting model for the weekly data. Figure 6.4 shows the ex ante and the smoothed regime probabilities in Panel (a), and the corresponding range observations in Panel (b). There is a clear peak in the smoothed and ex ante probabilities around the 1987 stock market crash. Also the high volatility period from 1997 until 2003 is clearly identified. Interestingly, the weekly data ranging back to the beginning of the 1960s also identify a longer period of high volatility from the end of the 1960 until the beginning of the 1980s. As already mentioned above, this period was characterized by many world economic changes and crises, as for example the first and the second oil shock and the collapse of the Bretton Woods system.

Figure 6.4: Regime probabilities (only the constant changes)



**Note:** In Panel (a) we show the ex ante (dotted line),  $p(S_t = 2|\Phi_{t-1})$ , and smoothed probabilities (solid line),  $p(S_t = 2|\Phi_T)$ , which are calculated as in Section 6.3.3 and 6.3.4 respectively. Both show the probability that the data at time  $t$  are drawn from the high volatility regime distribution. Panel (b) shows the corresponding observed range data for which the probabilities are calculated. All probabilities are obtained from the daily MS(2)-ARMA-C(1,1)-GARCH(4,0) model, where the C stands for only constant, meaning that only the constant is allowed to change with the regime.

In sum, our proposed model for the weekly log-range S&P500 data do a good job in terms of identifying important periods of financial uncertainty and increased

volatility in a very important US stock market index. They are capable of distinguishing quite clearly low- and high-volatility periods from each other. Also standardized residuals do not show important signs of autocorrelation or remaining unexplained conditional heteroscedasticity, which justifies the i.i.d. assumption important for quasi maximum likelihood estimation.

#### **Allowing all mean equation parameters to change**

Up to now we only allowed for changes in the constant term of the mean equation in Eq.(6.10). We also would like to check the evidence for changes in the dynamics. It might be that the dynamics of the range as a time series change with the regime. One might argue that in a high volatility regime the dependence of the volatility today on the volatility in the past is different compared to the low volatility regime because investors could change their behavior according to their perception of what volatility regime markets are in. In order to check for differences in the dynamics across regimes we let all parameters of the mean equation free to change with the regime. A note of caution here is that there might be other Markov-chains governing the switches in the intercept, the AR and MA parameters. If the true data generating process is governed by more than one Markov-chain, the estimation results and classification of regimes would be distorted. That is also why we present Sections 6.4.2 and 6.4.2 separately. In the estimation we follow the same approach as in the case where we only let the constant,  $\mu$ , change for identifying the appropriate model.

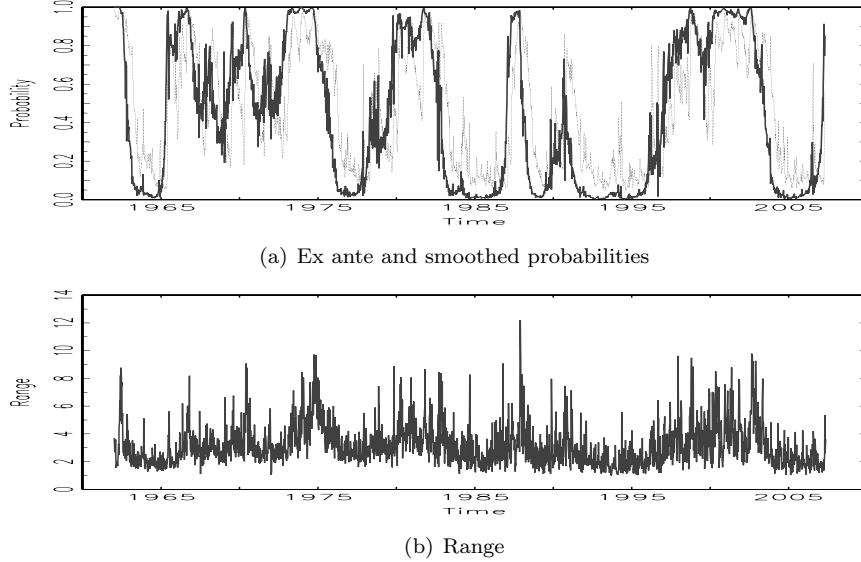
We present the estimation results in Table 6.3 and the ex ante and smoothed probabilities of the best fitting model in the corresponding Figure 6.5. Again we take the same approach for model selection as before. The best fitting model here, where we allow all parameters of the mean equation (6.10) to change, is the ARMA-X(1,1)-GARCH(2,1) specification.

Table 6.3: Results for weekly S&amp;P500 data allowing all mean equation parameters to change

Parameters	AM(1,0)		AM(1,1)		AM(1,1)-G(1,0)		AM(1,1)-G(1,1)		AM(1,1)-G(2,1)	
	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.
$\mu_1$	0.6000	0.0050	0.0288	0.0483	0.0281	0.0536	0.0227	0.0002	0.0244	0.0012
$\mu_2$	0.7005	0.0723	0.1711	0.0066	0.1555	0.3773	0.1357	0.0013	0.1103	0.0015
$a_{1,1}$	0.2326	0.0174	0.9655	0.0617	0.9651	0.0493	0.9745	0.0003	0.9670	0.0017
$a_{2,1}$	0.4550	0.0361	0.8714	0.0004	0.8809	0.2548	0.8969	0.0008	0.9132	0.0011
$b_{1,1}$			-0.8223	0.0546	-0.8347	0.1628	-0.8255	0.0009	-0.8388	0.0014
$b_{2,1}$			-0.5257	0.1408	-0.5537	0.3383	-0.4987	0.0005	-0.5838	0.0077
$\omega$	0.1130	0.0033	0.1055	0.0006	0.0957	0.0015	0.0002	0.0000	0.0001	0.0000
$\alpha_1$					0.0938	0.0302	0.0110	0.0000	0.0798	0.0103
$\alpha_2$									-0.0710	0.0104
$\beta_1$							0.9866	0.0002	0.9901	0.0048
$P_{11}$	0.9914	0.0033	0.9932	0.0020	0.9931	0.0149	0.9927	0.0000	0.9908	0.0000
$P_{22}$	0.9903	0.0031	0.9869	0.0125	0.9887	0.0117	0.9810	0.0001	0.9922	0.0005
Log-Likelihood	-828.018		-716.746		-708.093		-685.380		-680.768	
P-Values										
$LB_1$	0.000		0.716		0.988		0.877		0.917	
$LB_5$	0.000		0.772		0.774		0.907		0.887	
$LB_{10}$	0.000		0.847		0.825		0.965		0.936	
$LB_1^2$	0.000		0.002		0.000		0.046		0.164	
$LB_5^2$	0.000		0.000		0.000		0.107		0.243	
$LB_{10}^2$	0.000		0.000		0.000		0.094		0.177	
Jarque-Bera	30.241		27.562		38.670		34.113		34.713	
P-value	0.001		0.000		0.000		0.000		0.000	

**Note:** AM(a,b)-G(p,q) is short for an ARMA(a,b)-GARCH(p,q) specification. Parameters are estimated using the GAUSS6.0 conditional optimization package (co) under the constraints of all ARMA and GARCH roots lying outside the unit circle. Additionally, we impose a positivity constraint for the variance and conditional variance. We apply the standard convergence criteria. The parameters are as in Eq.(6.10) and (6.11) for the respective model specifications.  $P_{11}$  and  $P_{22}$  are the Markov chain transition probabilities for every period for staying in the low and in the high volatility regime, respectively.  $LB_x$  stands for the Ljung-Box test at  $x$  lags of the standardized residuals.  $LB_x^2$  is the same but for squared standardized residuals. For the Ljung-Box test we only report p-values for the null hypothesis of no autocorrelation. Ljung-Box test statistics are not adjusted for ARCH effects as suggested in Diebold (1986). The Jarque-Bera test tests for standard normality in the standardized residuals. For the Jarque-Bera test we report the test statistics and corresponding p-values.

Figure 6.5: Regime probabilities (all mean equation parameters change)



**Note:** In Panel (a) we show the ex ante (dotted line),  $p(S_t = 2|\Phi_{t-1})$ , and smoothed probabilities (solid line),  $p(S_t = 2|\Phi_T)$ , which are calculated as in Section 6.3.3 and 6.3.4 respectively. Both show the probability that the data at time  $t$  are drawn from the high volatility regime distribution. Panel (b) shows the corresponding observed range data for which the probabilities are calculated. All probabilities are obtained from the daily MS(2)-ARMA-X(1,1)-GARCH(1,1) model, where the X stands for all, meaning that all parameters in the ARMA equation are allowed to change with the regime.

In Table 6.3 three interesting results, compared to the earlier results where we only allowed the constant in the mean equation to change, appear. First, there seems to be a quite clear difference in the autoregressive coefficients across the regimes. In the case of the ARMA-X(1,1)-GARCH(2,1) specification we estimate the AR-coefficients for the low- and the high volatility regime to be equal to 0.9670 and 0.9132, respectively. This means that the half-life of a shock to the volatility is 21 weeks in the case of the low and 7 weeks for the high-volatility regime. So, in the low-volatility even 21 weeks after a shock around 50% of it is still present in the actual volatility. In the high-volatility regime markets seem to “forget” much more quickly. Here the half-life of a shock is around seven weeks. This confirms the results of Gray (1996) and Klaassen (2002) who also find that volatility persistence in the high volatility regime is lower. A second result is that also the moving-average parameter in the high-volatility regime are lower in absolute value than those in the low-volatility regime. The third interesting results is that the GARCH structure, being a GARCH(2,1), does not change compared to the results before and thereby appears to be very robust through different regimes



and specifications. So, also when we allow all parameters of the mean equation to change with the regime, there still is strong evidence for quite persistent volatility of the volatility.

We also experimented with specifications allowing all parameters of Eq.(6.10) and (6.11) to vary with the state. Results were inconclusive though, which probably is due to large amounts of coefficients that need to be estimated.<sup>14</sup> Another possibility is that the variance equation might be governed by a second Markov chain that not necessarily coincides with the Markov chain governing the parameters of the mean equation. We therefore, do not report those results here.

### 6.4.3 Comovements of volatility across asset markets

Above we showed how our Markov-switching model can be applied to the stock market data of the US. In the light of the foregoing chapters of this thesis it is clearly interesting to also have a look at other (inter)national asset markets in order to compare the results. For investors, for example, it is important to know if different asset markets tend to be in the same volatility regime at the same time because this has repercussions on hedging and portfolio diversification effects. We therefore apply the basic model to different asset markets in order to find possible comovements in the volatility regimes across countries and assets. Although this analysis will be limited in scope (for example, we limit ourselves to the most simple model specifications) we find evidence for interesting changes of volatility comovement in the sample.

The assets we are going to analyze are three months US T-bills, ten years US T-notes both starting in October 1965, the FTSE100 UK stock market index starting in April 1984, and the Japanese stock market index NIKKEI225 starting in August 1986. All sampled series run until February 2008. In the case of ten year T-notes we split the sample into two subsamples and fit each one separately to the model. The split is chosen to be in August 1979 the month when Paul Volcker became the Fed's chairman. In Chapter 5 of this thesis we already saw that at that point returns of government bonds started to behave differently than before and we found a structural break in the tail behavior of its return distribution. Also in Figure 6.8 one can see a clear and lasting change in the behavior of the range of T-notes. It is well known that Volcker ended the Fed's policy of targeting the interest rate but rather focused on the money supply by limiting the money growth. We take this as sufficient evidence for an exogenous change in the policy regime, that is also supported by our results in Chapter 5.

In the estimation procedure we follow exactly the same setup as before in the

---

<sup>14</sup>Results are available upon request.

S&P500 case. We choose that model as the best fit, which delivers standardized and squared standardized residuals that do not show any sign of remaining autocorrelation. In Table 6.4 we present the results of fitting our Markov-switching model to the different assets. We focus here on the model where we only allow the intercept to change with the regime because we only want to draw attention on the level of the volatility and not on changes in dynamics. As already explained in Section 6.3, also allowing the other parameters to change with the regime may distort the separation of regimes when the level of the volatility and its dynamics are governed by different Markov chains. In order to prevent such effects we limit ourselves on the level of volatility only. The structure of the table is similar to Tables 6.2 and 6.3 before. Here we only neglect the test statistics on the autocorrelations. The Jarque-Bera test statistics for normality for T-bills, T-notes, FTSE and Nickei cannot reject the null hypothesis at a 1% level. Results for the S&P500 stay as before and are repeated for convenience. Parameter estimates concerning the mean equations do not show strong qualitative differences across markets. S&P500, T-bills and the T-notes sample starting in 1979 clearly show significant GARCH effects. In the case of the other assets an ARCH(1) model is sufficient to account for the conditional variance of the range. Further, the Markov transition probabilities indicate that both the low and the high volatility regimes are persistent in all the cases.

Table 6.4: Estimation results for different weekly asset markets

Parameters	S&P500	TB	TN(79)	TN(08)	FTSE100	NIKKEI225
$\mu_1$	0.0545	0.1024	-0.0683	0.0210	0.1487	0.1461
$\mu_2$	0.0876	0.2452	0.0429	0.0641	0.2534	0.2350
$a_1$	0.9280	0.8049	0.7895	0.9479	0.8141	0.8355
$b_1$	-0.6907	-0.5514	-0.6530	-0.7863	-0.5554	-0.5743
$\omega$	0.0001	0.0087	0.3683	0.0001	0.1337	0.1430
$\alpha_1$	0.0892	0.0261	0.0992	0.0136	0.0270	0.0453
$\alpha_2$	-0.0797					
$\beta_1$	0.9893	0.9502		0.9860		
$P_{11}$	0.9961	0.9936	0.9879	0.9984	0.9953	0.9846
$P_{22}$	0.9962	0.9899	0.9853	0.9987	0.9904	0.9926
Log-Likelihood	-688.28	-2034.15	-708.52	-932.92	-547.03	-537.36
Jarque-Bera	32.86	2.98	7.97	2.95	5.90	5.89
P-value	0.00	22.55	1.87	22.89	5.24	5.24

**Note:** TB stands for T-bills. TN(79) and TN(08) refer to T-notes samples from 1962 until 1979 and from 1979 until 2008, respectively. Parameters are estimated using the GAUSS6.0 conditional optimization package (co) under the constraints of all ARMA and GARCH roots lying outside the unit circle. Additionally, we impose a positivity constraint for the variance and conditional variance. We apply the standard convergence criteria. The parameters are as in Eq.(6.10) for the respective model specifications.  $P_{11}$  and  $P_{22}$  are the Markov-chain transition probabilities for every period for staying in the low and in the high volatility regime, respectively. The Jarque-Bera test tests for standard normality in the standardized residuals. For the Jarque-Bera test we report the test statistics and corresponding p-values.

In Figures 6.6 to 6.10 we report the models' estimates for the regime probabilities and the underlying range data. The figures have again the same interpretation as Figures 6.4 and 6.5. Figure 6.6 just repeats the results of Figure 6.4 for convenience. At first sight there are obvious comovements of volatility regimes across asset markets. For example, all markets are found to be in a high volatility regime in the end of 1990 and in the beginning of the new millennium. Further we find that the US stock market, T-bills and T-notes markets are in the high volatility regime around the first and the second oil crises and the collapse of the Bretton Woods system.

In order to quantify the degree of comovements in the volatility regimes across asset markets we calculate concordance indices for different asset combinations. The concordance index can be calculated as follows:

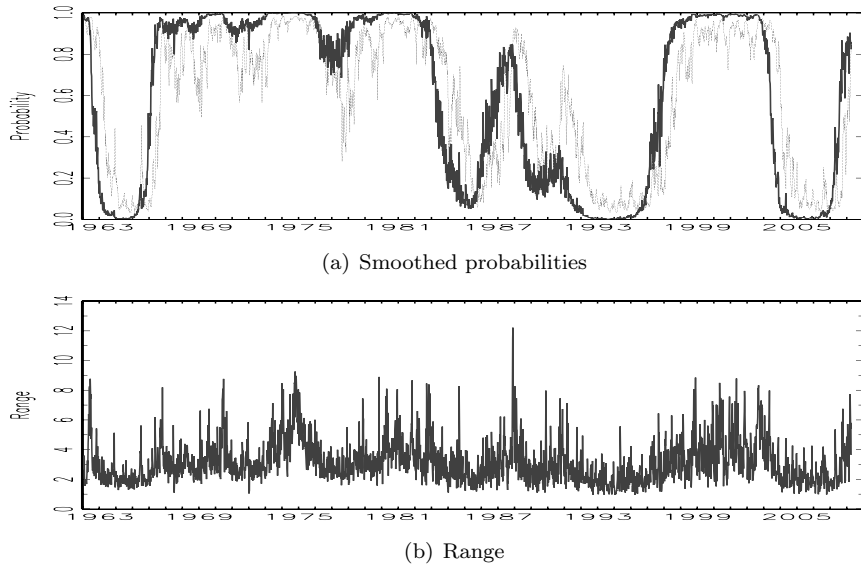
$$CI_{i,j} = \frac{1}{T} \sum_{t=1}^T [H_t^i H_t^j + L_t^i L_t^j], \quad (6.27)$$

where  $H_t^i$  is an indicator function that equals one if market  $i$  is in the high volatility regime at time  $t$  and  $L_t^i$  is equal to one if it is in the low volatility regime. Thereby, the index can only assume values between 0 and 1. It equals 0 if there is absolutely no concordance and it equals 1 if there is perfect concordance. The extension to the case with more than two markets is straightforward. In order to distinguish low and high volatility regimes we use the estimated smoothed probabilities and follow the rule that  $H_t^i = 1$  when  $p(S_t|\Phi_T) > 0.5$ , where  $p(S_t|\Phi_T)$  is the probability that the observation of the range at time  $t$  came from the high volatility regime. When that probability is smaller than 0.5 we classify it as coming from the low volatility regime. Other classification algorithms are obviously possible, but the cutoff at 50% appears to be the most natural one.

In Figure 6.11 we show the results of recursive concordance index calculations. This means that at time  $t$  the value of the concordance index reflects the concordance of the sample up to and including point  $t$ . Therefore, the last observation then gives the concordance of the considered assets for the full sample. Like this we can see the development of the degree of comovement of assets' volatilities over time but also keep the overall comparability over time. By construction the recursive concordance index shows larger variability at the beginning than at the end of the sample.

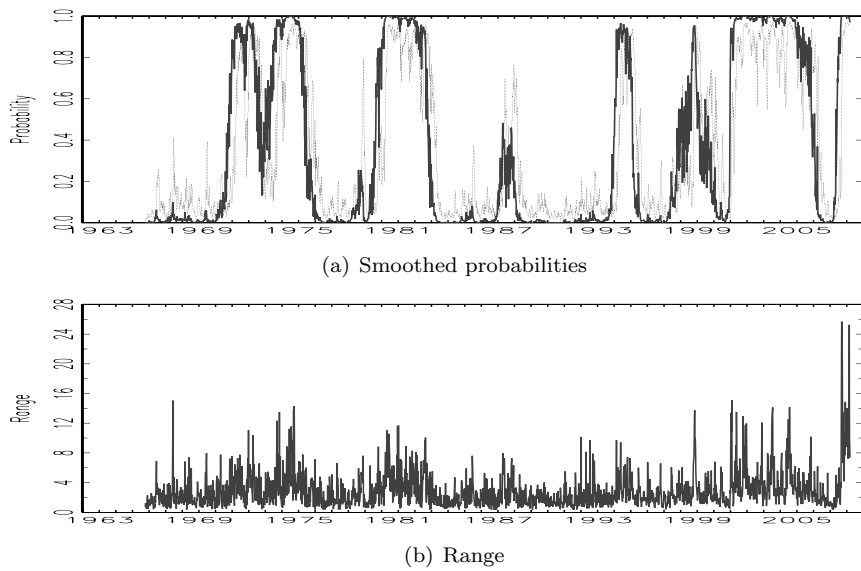
In general, over time the figures show a tendency of increasing comovements of the volatility regimes across assets and countries. Especially the S&P500 with the FTSE100 and T-notes with T-bills clearly show such a tendency. Here also S&P500 and FTSE100 show very high absolute values of concordance even reaching higher than 0.9 at the end of the sample period. This clearly shows a strong comovement

Figure 6.6: S&amp;P500 range and smoothed probabilities



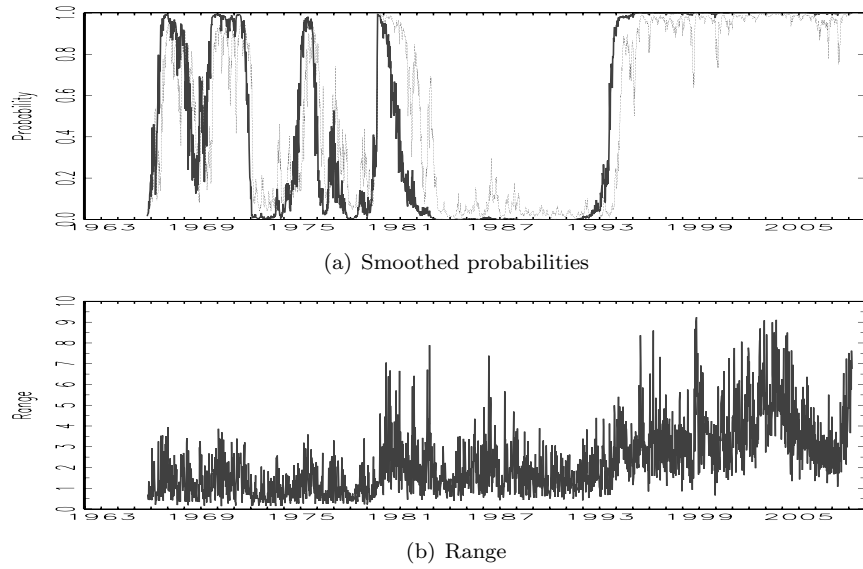
**Note:** Range is equal to  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ .

Figure 6.7: T-bills range and smoothed probabilities



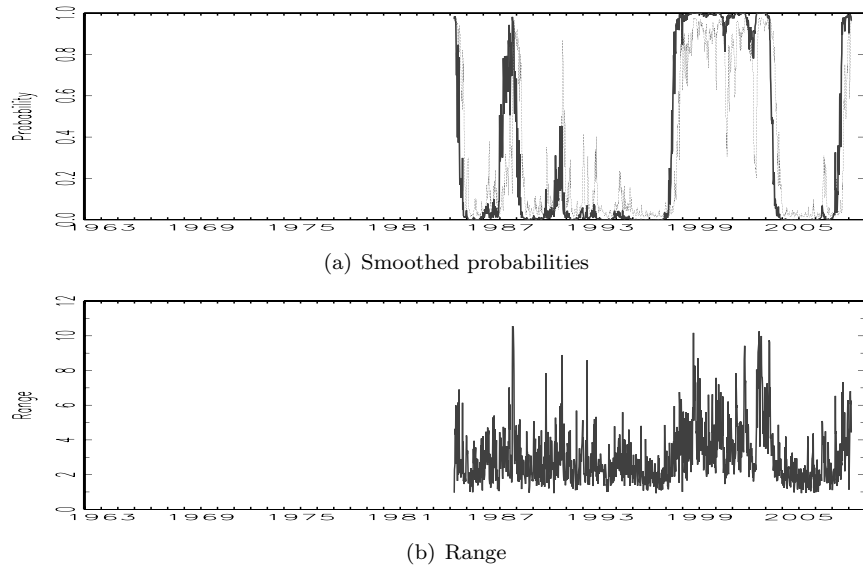
**Note:** Range is equal to  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ .

Figure 6.8: T-notes range and smoothed probabilities



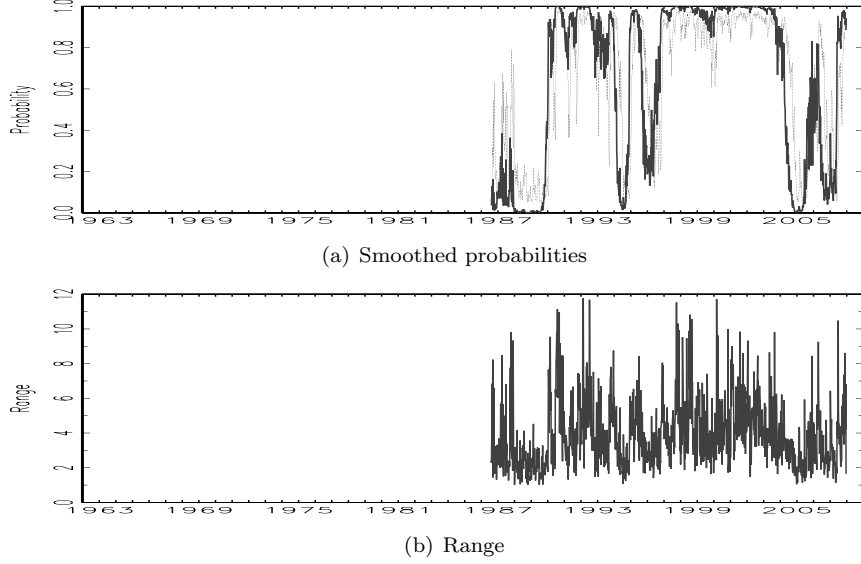
**Note:** Range is equal to  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ .

Figure 6.9: FTSE range and smoothed probabilities



**Note:** Range is equal to  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ .

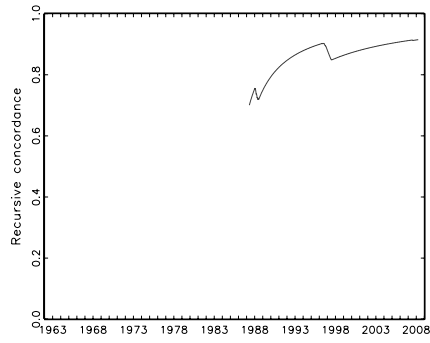
Figure 6.10: NIKKEI range and smoothed probabilities



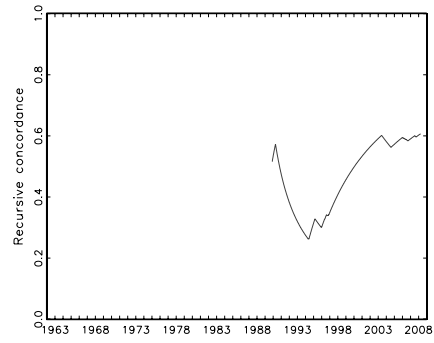
**Note:** Range is equal to  $R_t = 100 * (p_t^{Max} - p_t^{Min})$ .

of volatility regimes of the US stock market with the UK stock market, whereas the comovement of the US with the Japanese stock market is much lower in absolute value. Another clear point is the strong increase in comovement of volatility regimes among the considered stock markets beginning in the mid and late 1990s. Such an increase in concordance is strongly driven by the high volatility regime starting at the end of the nineties and reaching far into the new millennium. This long period of high volatility we find in virtually all stock markets and seems to be an international or even global phenomenon. A further interesting observation is also the clear increase in comovement of volatility regimes among the US stock, T-notes, and T-bills markets. It appears that also those markets show a tendency towards stronger co-volatility and thereby increasing integration through time. Such an tendency of increasing comovements of volatility across asset classes and asset markets can either be due to some global factors affecting all markets similarly or it reflects a stronger international and inter-asset integration of the markets. An identification of the precise causes of the increased comovements is not possible with these methods and beyond the scope of this chapter.

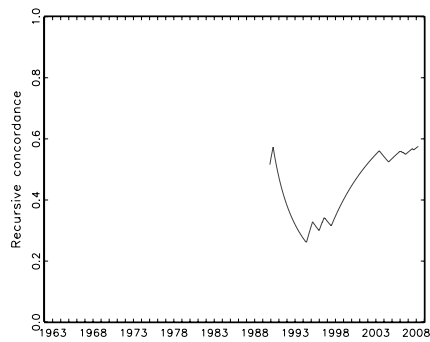
Figure 6.11: Concordances



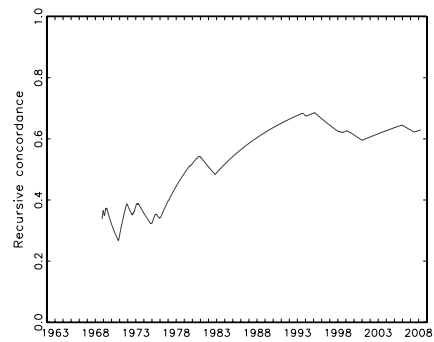
(a) S&P500,FTSE



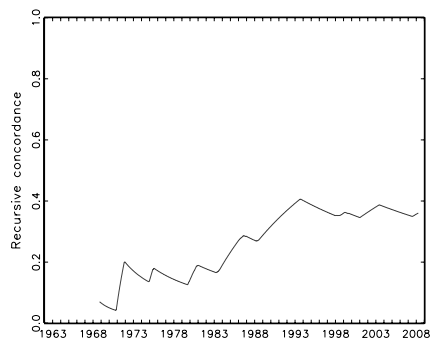
(b) S&P500,NIKKEI



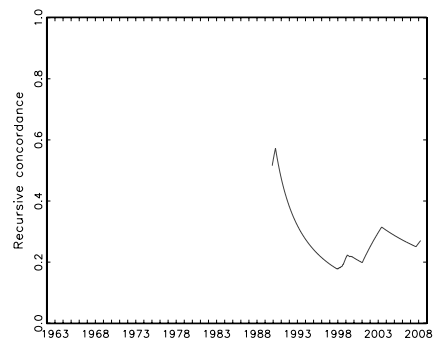
(c) S&P500,FTSE,NIKKEI



(d) T-bills,T-notes



(e) S&P500,T-bills,T-notes



(f) S&P500,T-bills,T-notes,FTSE,NIKKEI

**Note:** Concordances are calculated using (6.27) recursively.

#### 6.4.4 Forecasting performance

A better fit to the data is already an own end for modeling the data generating process of volatility as a Markov-switching model instead of a linear one in order to identify high and low volatility periods within the sample. But another interesting point is a comparison of forecasting performances. In this subsection we present the results of an in-sample forecasting comparison of our proposed Markov-switching model with a linear ARMA-GARCH specification. We estimate the best fitting linear<sup>15</sup> and the best fitting MS model using the full sample. Then we pick a starting point  $t$  in the sample and forecast  $F$  periods into the future. After obtaining such a forecast we go to observation  $t + 1$  and do the same again rolling through the sample until we arrive at period  $T - F$  which is the period of the last forecast. An underlying assumption of such a procedure is that the parameter estimates do not change much by either estimating the models with the full sample or by always re-estimating it.<sup>16</sup> For calculating our forecasts we follow the methods developed in Davidson (2004) where he proposes a method for multi-period forecasting with a Markov-switching dynamic regression model accounting for conditional heteroscedasticity.

Imagine that we want to forecast  $r_{t+F}$  for  $F \geq 1$  given observations on the process up to date  $t$ . With other words the object of interest is  $E(r_{t+F}|\Phi_t)$ . Davidson (2004) develops a recursion for computing  $E(r_{t+F}|\Phi_t)$ , which we denote by  $\hat{r}_{t+F}$  for brevity. Such a recursion involves only  $K$  terms at each iteration. The terms are the probability-weighted averages of the one-step contingent forecasts. We can rewrite Davidson's recursion slightly by adapting it to our case as:

$$\hat{r}_{t+F} = \sum_{j_F=1}^K \hat{P}_{F,j_F} [\mu_{j_F} + \sum_{f=1}^a a_{m,j_F} \hat{r}_{F-f} + \sum_{f=F}^{F+a} a_{m,j_F} \hat{r}_{F-f}] \quad (6.28)$$

$$+ \sum_{f=\min(b,F)}^b b_{m,j_F} \hat{e}_{F-f}, \quad (6.29)$$

where  $\hat{P}_{f,j} = Pr(S_f = j|\Phi_t)$  and is generated from

$$\hat{P}_{f,j} = \sum_{i=1}^K P_{ji} \hat{P}_{f-1,i}, \quad \text{for } j = 1, \dots, K \quad \text{and } f = 1, 2, \dots, F.$$

<sup>15</sup>The linear model we consider is an ARMA(1,1)-GARCH(2,1) specification without Markov-switching. For brevity we do not show the details of this model here, but they are available upon request.

<sup>16</sup>We performed estimations of the models only using sub-samples. It turned out that such an assumption appears to be justified.



For a proof see Davidson (2004, p.3-4).

As a penalty function we use the mean absolute and the mean squared errors. We apply Eq.(6.28) to different forecasting horizons  $F = 1, 5, 10, 20, 25$  and show their relative performances with respect to the linear model in Table 6.5.

Table 6.5: Point forecast comparison

Panel A		Only constant changes			
Criterion	1	5	10	20	25
Absolute	0.611	2.140	3.379	4.367	3.932
Squared	1.119	3.555	5.932	6.996	6.893
$S_1$	0.021	0.002	0.006	0.038	0.099
$S_1^{adj}$	0.021	0.002	0.007	0.044	0.110
$S_2$	0.045	0.001	0.000	0.000	0.001
$S_{2a}$	0.049	0.001	0.000	0.000	0.001
$S_{3a}$	0.025	0.000	0.000	0.000	0.000
Panel B		All mean equation parameters change			
Absolute	-26.524	-9.420	-3.215	1.528	2.974
Squared	-60.996	-21.229	-8.390	3.347	5.644
$S_1$	1.000	1.000	0.922	0.287	0.185
$S_1^{adj}$	1.000	1.000	0.918	0.294	0.197
$S_2$	1.000	0.999	0.649	0.009	0.097
$S_{2a}$	1.000	0.999	0.665	0.010	0.105
$S_{3a}$	1.000	0.999	0.929	0.087	0.013

**Note:** Values are the improvements in the forecasting performance of the Markov-switching model compared to the linear model. We compare the ARMA(1,1)-GARCH(2,1) linear model with the MS-ARMA(1,1)-C-GARCH(2,1) model starting the forecasts at  $t = 1800$  which corresponds to ca. the 75<sup>th</sup> percentile of the sample. All test statistics but  $S_{1adj}$  are as in Diebold and Mariano (1995).  $S_1^{adj}$  is the adjusted  $S_1$  test statistic from Diebold and Mariano (1995) as suggested in Harvey et al. (1997). Test statistics are based on the absolute forecasting error.

In order to compare the forecasting accuracy of the linear and the Markov-switching model more formally we perform the statistical tests proposed in Diebold and Mariano (1995). They develop different test statistic allowing to compare forecasts from two competing models against each other. Such a comparison is based on a loss function  $g(\cdot)$  that can take a variety of forms. We opt for an absolute forecasting error loss function. The Diebold and Mariano procedure tests the null hypothesis of equality of the two competing forecasts against the alternative that one forecasting model outperforms the other in its forecasting accuracy. In equation form the null hypothesis may be written as:

$$E[g(e_{it})] = E[g(e_{jt})], \quad \text{or} \quad E(d_t) = 0,$$

where  $e_{it}$  is model  $i$ 's forecasting error and  $d_t \equiv [g(e_{it}) - g(e_{jt})]$  is the loss differential. They propose different test statistics one of them being an asymptotic

test, which they call  $S_1$ :

$$S_1 = \frac{\bar{d}}{\sqrt{\widehat{Var}_{\bar{d}}/T_f}},$$

where

$$\bar{d} = \frac{1}{T_f} \sum_{f=1}^{T_f} d_t \quad (6.30)$$

$$Var_{\bar{d}} = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j, \quad \gamma_j = cov(d_t, d_{t-j}) \quad (6.31)$$

and where  $\widehat{Var}_{\bar{d}}$  is a consistent estimate of the asymptotic variance of  $\sqrt{T}\bar{d}$  as proposed in Diebold and Mariano. The infinite sum of covariances in Equation (6.30) is difficult to estimate. (Diebold and Mariano, 1995, p.254) state that “optimal k-step ahead forecast errors are at most (k-1)-dependent...(k-1)-dependence implies that only (k-1) sample autocovariances need to be used...” They further show that  $S_1 \stackrel{a}{\sim} N(0, 1)$ .<sup>17</sup>

One might argue if an asymptotic test is applicable to our data. So, we also calculate the “finite-sample tests” proposed by (Diebold and Mariano, 1995), which we will show below:

$$S_2 = \sum_{t=1}^T I_+(d_t),$$

where

$$I_+(d_t) = 1 \quad \text{if } d_t > 1 \\ = 0 \quad \text{otherwise.}$$

$S_2$  may be assessed using the cumulative binomial distribution with a success probability of  $p = 0.5$  under the null. In large samples another version of the  $S_2$  sign test is:

$$S_{2a} = \frac{S_2 - 0.5T}{\sqrt{0.25T}} \stackrel{a}{\sim} N(0, 1).$$

The last test statistic we will use is based on a rank-test and is also standard

---

<sup>17</sup>We also report the test statistic  $S_1^{adj} = \left[ \frac{T_f + 1 - 2F + T_f^{-1}F(F-1)}{T_f} \right]^{1/2} S_1$  as suggested in Harvey et al. (1997) to account for finite sample bias and heavy-tailed error distributions. Results do not differ much from  $S_1$ .

normally distributed under the null:

$$S_{3a} = \frac{S_3 - \frac{T(T-1)}{4}}{\sqrt{\frac{T(T+1)(2T+1)}{24}}} \stackrel{a}{\sim} N(0, 1),$$

where

$$S_3 = \sum_{t=1}^T I_+(d_t) \text{rank}(|d_t|).$$

Again in Table 6.5 we show the results of forecasting comparison between the linear and the non-linear Markov-switching model. It is apparent that in the weekly dataset the Markov-switching model, where only the constant term in the mean equation changes, outperforms the linear alternative significantly at any forecasting horizon considered. It is interesting but not surprising to see that the forecasting accuracy of the Markov-switching compared to the linear model improves the longer the forecasting horizon. Such a behavior was to be expected because the change in the absolute difference in the intercepts between the low and the high volatility state is not large and the processes need some time after state-switches to “burn-in” towards the new unconditional volatility level.

When the Markov-switching model, where all mean parameters are free to change, is the competing one, we can see that the linear model forecasts better at short horizons and marginally worse at longer horizons. The better forecasting performance of the MS model at longer horizons is at most small and not very significant. There are a couple of possible explanations for such an outcome. One explanation might be that the differences between the linear and the non-linear MS model are not very large not leading to any significant improvements. Another reason can be that the process does not remain long enough in one regime or another in order to take full advantage of the difference in constants across regimes. This would not allow the forecast to burn in towards the respective unconditional mean in order to obtain a better forecast performance. Such a reason might be justified by again having a look at Figure 6.5, where it is apparent that the average time the process is estimated to stay in one of the two regimes is much shorter than for the MS model which only allows for changes in the intercept of the mean equation.

## 6.5 Conclusions

In this chapter we propose a new non-linear volatility model based upon the log-range being defined as the log of the spread (the range) between the observed

maximum and the minimum price of assets within a trading week. The results of such an analysis are of potential interest for option pricing, hedging decisions, VaR calculations, but also for policy making. We find strong evidence for an underlying and unobservable Markov chain governing the parameters of the ARMA-GARCH specification that fits the log-range data best. We clearly identify two, a high and a low, volatility regimes. Smoothed regime probabilities that are obtained during the estimation of the models also very well coincide with periods of either low or high volatility observed in the data. Periods most likely to show stronger than average volatility correspond to the collapse of the Bretton Woods system, the first and the second oil crisis, to a (surprisingly) lower extend the period around “Black Monday” in October 1987, and the time from 1998 until 2003 with the Russian crisis, the burst of the dot-com bubble and the second Gulf War.

We further find evidence for different volatility dynamics across different volatility regimes. Volatility appears to be more persistent when the average level of it is relatively low, but seems to be less persistent when it is high. Such results confirm those of Gray (1996) and Klaassen (2002) and hint at the fact that asset market participants act differently during normal versus very volatile periods. In high volatility periods they seem to “forget” quicker than during low volatility periods. Such results have to be interpreted with caution, though, as the dynamics can be governed by another Markov chain as the change in the intercept.

The conditional volatility of the log-range (or the volatility of the volatility) is found to be described well by a GARCH structure with strong persistence, which is very robust over all different models considered. Such a fact means that shocks to the volatility of the volatility in the S&P500 stock index tend to be still present in the market many periods after they happened.

A comparison across different asset markets revealed pronounced comovements in the volatility regimes, especially among stock markets like the S&P500 and the FTSE100. These concordances in volatility also show an increasing tendency over the sample period suggesting the conclusion that there are either more pronounced global factors affecting all markets at the same time or that asset markets have become more integrated. Obviously this does not preclude a combination of both.

A forecasting comparison between a linear model and the proposed Markov-switching models shows promising results. Whereas the Markov-switching model allowing all mean equation parameters to change performs only marginally better at longer horizons than the linear model, we find that the Markov-switching specification only allowing the constant term to change with the regime performs significantly better than the linear competitor at all horizons considered.

Much remains to be done in the area of volatility estimation and forecasting. Our model combining nonparametric volatility estimation with parametric

Markov-switching time series methods is not the end of the story. Some very interesting extensions of our model might include the possibility of more than two volatility regimes. The transition probabilities between regimes do not need to be constant either, but can be specified to be dependent on exogenous variables. We might also want to allow for more than one Markov chain governing different aspects of the model. Another very interesting extension of our model would be to check if the forecasting performance of the Markov-switching model may be improved by assuming that the unconditional mean of the volatility changes with regimes and not only the constant term. Such a behavior would cause the forecasts to move much quicker to the new mean of the volatility corresponding to the respective regime the process is forecast to be in. We are working on some of these extensions and it will be interesting to see to what extent they might improve the estimation and forecasting of asset market volatility.

---

## Bibliography

- Abraham, A. and D. Ikenberry (1994). The individual investor and the weekend effect. *Journal of Financial and Quantitative Analysis* 29, 263–277.
- Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance* 57(3), 1047–1091.
- Andersen, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long run in high frequency returns. *Journal of Finance* 52, 975–1005.
- Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885–905.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2006). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Review of Economics and Statistics*, forthcoming.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics* 61(1), 43–76.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96, 42–55.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 71, 579–625.

## BIBLIOGRAPHY

---

- Andersen, T. G., T. Bollerslev, and X. Huang (2007). A semiparametric framework for modelling and forecasting jumps and volatility in speculative prices. Working paper, Duke University.
- Anderson, D. C., A. F. Darrat, and M. Zhong (2003). Do us stock prices deviate from their fundamental values? *Journal of Banking and Finance* 27, 673–697.
- Anderson, H. M. and F. Vahid (2007). Forecasting the volatility of australian stock returns: Do common factors help? *Journal of Business & Economic Statistics* 25, 76–90. American Statistical Association.
- Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). John Wiley and Sons.
- Andrews, D. W. K. (1993). Tests for parameter stability and structural change with unknown change point. *Econometrica* 59, 817–858.
- Andrews, D. W. K. and M. Buchinsky (2000). A three-step method for choosing the number of bootstrap repetitions. *Econometrica* 68(1), 23–51.
- Andrews, D. W. K. and R. C. Fair (1988). Inference in nonlinear econometric models with structural change. *The Review of Economic Studies* 55, 615–639.
- Andrews, D. W. K. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62(2), 1383–1414.
- Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70(1), 191–221.
- Baillie, R. T. and T. Bollerslev (1989). The message in daily exchange rates: A conditional variance tale. *Journal of Business and Economic Statistics* 7, 297–305.
- Balassa, B. (1961). *The theory of economic integration*. London, Allen & Unwin.
- Ball, C. A. and W. N. Torous (1984). The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. *Journal of Business* 57, 97–112.
- Barndorff-Nielsen, O. E. and N. Shepard (2001). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* 64, 253–280.

- Barndorff-Nielsen, O. E. and N. Shepard (2003). Realized power variation and stochastic volatility. *Bernoulli* 9, 243–265. (correction, pp.1109–1111).
- Barndorff-Nielsen, O. E. and N. Shepard (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1–48.
- Barndorff-Nielsen, O. E. and N. Shepard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1–30.
- Beckers, S. (1983). Variance of security price returns based on high, low, and closing prices. *Journal of Business* 56, 97–112.
- Beine, M., J. Lahaye, S. Laurent, C. Neely, and F. C. Palm (2007). Central bank intervention and exchange rate volatility, its continuous and jump components. *forthcoming in International Journal of Finance and Economics*.
- Beine, M., J. Lahaye, C. J. Neely, F. C. Palm, and S. Laurent (2007). Central bank intervention and exchange rate volatility, its continuous and jump components. *International Journal of Finance and Economics* 12(2), 201–223.
- Beirlant, J., G. Dierckx, Y. Goegebeur, and G. Matthys (1999). Tail index estimation and the exponential regression model. *Extremes* 2(2), 177–200.
- Bekaert, G., C. Harvey, and A. Ng (2005). Market integration and contagion. *Journal of Business* 78(1), 39–70.
- Bekaert, G. and C. R. Harvey (1995). Time-varying world market integration. *Journal of Finance* 50(2), 403–444.
- Bekaert, G. and C. R. Harvey (2000). Foreign speculators and emerging equity markets. *Journal of Finance* 55(2), 565–613.
- Bekaert, G. and G. Wu (2000). Asymmetric volatility and risk in equity markets. *The Review of Financial Studies* 13, 1–42.
- Bhattacharya, R. N., V. K. Gupta, and E. Waymire (1983). The hurst effect under trends. *Journal of Applied Probability* 20, 649–662.
- Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.
- Bollen, N. P. B., S. F. Gray, and R. E. Whaley (2000). Regime switching in foreign exchange rates: Evidence from currency option prices. *Journal of Econometrics* 94, 239–276.



## BIBLIOGRAPHY

---

- Bollerslev, T. (1986). Generalised autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rate: A multivariate generalized arch approach. *Review of Economics and Statistics* 72, 498–505.
- Bollerslev, T., J. Litvinova, and G. Tauchen (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* 4(3), 353–384.
- Brandt, M. W. and F. X. Diebold (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business* 79(1), 61–73.
- Bry, G. and C. Boschan (1971). Cyclical analysis of time series: Selected procedures and computer programs. *NBER*. New York.
- Burns, A. M. and W. C. Mitchell (1946). *Measuring business cycles*. New York: National Bureau of Economic Research.
- Cai, J. (1994). A markov model of unconditional variance in arch. *Journal of Business and Economic Statistics* 12, 309–316.
- Calvo, S. and C. M. Reinhard (1996). Capital flows to latin america: Is there evidence of contagion effects? In G. Calvo, M. Goldstein, and E. Hochreiter (Eds.), *Private Capital Flows to Emerging Markets*. Institute for International Economics.
- Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The econometrics of financial markets*. Princeton: Princeton University Press.
- Candelon, B. and V. Bodart (2004). Evidences of contagion and interdependence using a frequency domain framework. Working paper, Maastricht University.
- Candelon, B., J. Piplack, and S. Straetmans (2008a). Multivariate business cycle synchronization in small samples. Working Paper.
- Candelon, B., J. Piplack, and S. Straetmans (2008b). On measuring synchronization of bulls and bears: The case of east asia. *Journal of Banking and Finance* 32(6), 1022–1035.
- Carlson, M. (2007). A brief history of the 1987 stock market crash with a discussion of the federal reserve response. Staff working paper 2007-13, Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

- Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica* 51, 1305–1324.
- Chauvet, M. and S. Potter (2000). Coincident and leading indicators of the stock market. *Journal of Empirical Finance* 7, 87–111.
- Chen, G. M., M. Firth, and O. Rui (2002). Stock market linkages: Evidence from latin america. *Journal of Banking and Finance* 26(6), 1113–1141.
- Chen, N.-f. and F. Zhang (1997). Correlations, trades and stock returns of the pacific rim markets. *Pacific Basin Finance Journal* 5, 559–577.
- Chen, Z. and P. Knez (1995). Measurement of market integration and arbitrage. *Review of Financial Studies* 8(2), 287–325.
- Cheung, Y.-W. (2007). An empirical model of daily highs and lows. *International Journal of Finance and Economics* 12, 1–20.
- Cheung, Y.-W. and U. Erlandsson (2005). Exchange rates and markov switching dynamics. *Journal of Business and Economic Statistics* 23(3), 314–320.
- Cheung, Y.-W. and L. K. Ng (1996). A causality-in-variance test and its application to financial market prices. *Journal of Econometrics* 72, 33–48.
- Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (carr) model. *Journal of Money Credit and Banking* 37, 561–582.
- Christiano, L. and W. den Haan (1996). Small sample properties of gmm for business cycle analysis. *Journal of Business and Economic Statistics* 14(3), 309–327.
- Claessens, S., R. Dornbusch, and Y. C. Park (2001). Contagion: Why crises spread and how this can be stopped. In S. Claessens and K. Forbes (Eds.), *International Financial Contagion*. Kluwer Academic Publishers.
- Click, R. W. and M. G. Plummer (2005). Stock market integration in asean after the asian financial crisis. *Journal of Asian Economics* 16, 5–28.
- Coakley, J. and A. M. Fuertes (2006). Valuation ratios and price deviations from fundamentals. *Journal of Banking and Finance* 30, 2325–2346.
- Cochrane, J. H. (2001). *Asset pricing*. Princeton: Princeton University Press.
- Cross, F. (1973). The behavior of stock prices on fridays and mondays. *Financial Analysts Journal* 29(6), 67–69.

## BIBLIOGRAPHY

---

- Davidson, J. (2004). Forecasting markov-switching dynamic, conditional heteroscedastic processes. Discussion paper, Cardiff University.
- Davidson, J. E. H. (1994). *Stochastic limit theory*. Oxford: Oxford University Press.
- Davidson, J. E. H. (2000). *Econometric theory*. Oxford: Blackwell Publishing.
- De Gregorio, J., S. Edwards, and R. O. Valdes (2000). Controls on capital inflows: Do they work? *Journal of Development Economics* 63(1), 59–83.
- De Haan, L., D. W. Jansen, K. Koedijk, and C. G. de Vries (1994). Safety first portfolio selection, extreme value theory and long run asset risks. In J. Galombos (Ed.), *Extreme Value Theory and Applications*, pp. 471–487. Dordrecht: Kluwer Academic Publishers.
- De Jong, R. M. and T. M. Woutersen (2007). Dynamic time series binary choice. Working paper.
- DeLong, J. B. (1992). Bull and bear in the united states. In P. Newman, M. Milgate, and J. Eatwell (Eds.), *The New Palgrave Dictionary of Money and Finance*. London: Press Limited.
- Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.
- Diebold, F. X. (1986). Testing for serial correlation in the presence of arch. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 323–328. Washington, DC: American Statistical Association, 1987.
- Diebold, F. X. and A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics* 105, 131–159.
- Diebold, F. X., J.-H. Lee, and G. C. Weinbach (1994). Regime switching with time-varying transition probabilities. In C. Hargreaves (Ed.), *Nonstationary Time Series Analysis and Cointegration: Advanced Texts in Econometrics*. Oxford and New York: Oxford University Press.
- Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13(3), 253–263.
- Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: A multivariate latent factor arch model. *Journal of Applied Econometrics* 4, 1–21.

- Draisma, G., H. Drees, A. Ferreira, and L. de Haan (2001). Tail dependence in independence. Report 2001-014, Eurandom, Institute for Statistics, Probability, Operations Research, and their Applications.
- Drees, H. (2003). Extreme quantile estimation for dependent data with applications to finance. *Bernoulli* 9(4), 617–657.
- Dunbar, N. (2001). *Inventing money: The story of Long-Term Capital Management and the legends behind it*. John Wiley & Sons.
- Dungey, M., R. Fry, and V. Martin (2006). Correlation, contagion and asian evidence. *Asian Economic Papers* 5(2), 32–72.
- Ederington, L. H. (1979). The hedging performance of the new futures market. *The Journal of Finance* 34(1), 157–170.
- Edwards, S., J. Gómez Biscarri, and H. F. Pérez de Gracia (2003). Stock market cycles, financial liberalization and volatility. *Journal of International Money and Finance* 22, 925–955.
- Edwards, S. and R. Rigobon (2002). Currency crises and contagion: An introduction. *Journal of Development Economics* 69(2), 307–313.
- Eichengreen, B., A. K. Rose, and C. Wyplosz (1996). Contagious currency crises. Working Paper 5681, National Bureau of Economic Research.
- Einstein, A. (1954). Science and religion. In D. Bronstein and H. Schulweis (Eds.), *Approaches to the philosophy of religion: A book of readings*, pp. 68–72. Ayer Publishing.
- Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling extremal events*. Berlin: Springer.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50, 987–1007.
- Engle, R. F. and C. W. J. Granger (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–276.
- Engle, R. F. and S. Kozicki (1993). Testing for common features. *Journal of Business and Economic Statistics* 11(4), 369–380.
- Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous garch. *Econometric Theory* 11, 122–150.

## BIBLIOGRAPHY

---

- Engle, R. F. and J. Marcucci (2006). A long-run pure variance common features model for the common volatilities of the dow jones. *Journal of Econometrics* 132(1), 7–42.
- Engle, R. F., V. K. Ng, and M. Rothschild (1990). Asset pricing with a factor arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics* 45, 213–238.
- Engle, R. F., V. K. Ng, and M. Rothschild (1994). Volatility and links between national stock markets. *Econometrica* 62, 901–934.
- Engle, R. F. and R. Susmel (1993). Common volatility in international equity markets. *Journal of Business and Economic Statistics* 11(2), 167–176.
- Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25, 383–417.
- Feller, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *The Annales of Mathematical Statistics* 22, 427–432.
- Fenton-O’Creevy, M., N. Nicholson, E. Soane, and P. Willman (2004). *Traders - risks, decisions, and management in financial markets*. Oxford University Press.
- Forbes, K. J. and R. Rigobon (2002). No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance* 57(5), 2223–2261.
- Garcia, R. (1998). Asymptotic null distribution of the likelihood ratio test in markov-switching models. *International Economic Review* 39, 763–788.
- Garman, M. B. and M. J. Klass (1980). On the estimation of price volatility from historical data. *Journal of Business* 53, 67–78.
- Gelos, R. G. and R. Sahay (2001). Financial market spillovers in transition economies. *Economies of Transition* 9(1), 53–86.
- Gerlach, S. and F. Smets (1995). Contagious speculative attacks. *European Journal of Political Economy* 11, 45–63.
- Ghosh, A., R. Saidi, and K. H. Johnson (1999). Who moves the asia-pacific stock markets: US or Japan? empirical evidence based on the theory of cointegration. *Financial review* 34, 159–170.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: Getting the most out of return data sampled at different frequencies. Working Paper.

- Giot, P., S. Laurent, and M. Petitjean (2007). Trading activity, realized volatility and jumps. Working paper.
- Gobbi, F. and C. Mancini (2006). Identifying the diffusion covariation and the co-jumps given discrete observations. Working paper.
- Goetzmann, W. N., L. Li, and K. G. Rouwenhorst (2005). Long-term global market correlations. *Journal of Business* 78(1), 1–38.
- Goldie, C. and R. Smith (1987). Slow variation with remainder: Theory and applications. *Quarterly Journal of Mathematics* 38, 45–71.
- Gómez Biscarri, J. and H. F. Pérez de Gracia (2003). Stock market cycles and stock market development in Spain. Working Paper.
- Gordin, M. I. (1969). The central limit theorem for stationary processes. *Soviet Mathematics-Doklady* 10, 1174–1176.
- Gourieroux, C., A. Holly, and A. Monfort (1982). Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. *Econometrica*, 63–80.
- Gray, S. F. (1995). An analysis of conditional regime-switching models. Working paper (Fuqua School of Business, Duke University, Durham, NC).
- Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* 42, 27–62.
- Haerdle, W., J. L. Horowitz, and J.-P. Kreiss (2003). Bootstrap methods for time series. *International Statistical Review* 71(2), 435–459.
- Hall, A. R. (2005). *Generalized Method of Moments*. New York: Oxford University Press.
- Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society B* 44(1), 37–42.
- Hall, P. and J. L. Horowitz (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica* 64(4), 891–916.
- Hall, P., J. L. Horowitz, and B. Jing (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* 82(3), 561–574.
- Hamao, Y., R. W. Masulis, and V. Ng (1990). Correlations in price changes and volatility across international stock markets. *Review of Financial Studies* 3, 281–307.

## BIBLIOGRAPHY

---

- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57(2), 357–384.
- Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics* 45, 39–70.
- Hamilton, J. D. and G. Lin (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics* 11, 573–593.
- Hamilton, J. D. and R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 64, 307–333.
- Hannan, B. (1980). *Binary time series*. New York: Marcel Dekker.
- Hannan, E. J. (1970). *Multiple time series*. New York: John Wiley and Sons.
- Hansen, B. E. (1992). The likelihood ratio test under non-standard conditions: Testing the markov trend model of gnp. *Journal of Applied Econometrics* 7, S61–S82. supplement.
- Hansen, B. E. (1996). Erratum: The likelihood ratio test under non-standard conditions: Testing the markov trend model of gnp. *Journal of Applied Econometrics* 11, 195–198.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50(4), 1029–1054.
- Harding, D. and A. R. Pagan (2006a). The econometric analysis of constructed binary time series. Technical Report 963, The University of Melbourne.
- Harding, D. and A. R. Pagan (2006b). Synchronization of cycles. *Journal of Econometrics* 132(1), 59–79.
- Hartmann, P., S. Straetmans, and C. de Vries (2004). Asset market linkages in crisis periods. *Review of Economics and Statistics* 86(1), 313–326.
- Hartmann, P., S. Straetmans, and C. de Vries (2005). Banking system stability: A cross-atlantic perspective. NBER Working Paper.
- Harvey, D., L. Leybourne, and P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13, 281–291.
- Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* 3(5), 1163–1173.

- Horowitz, J. (2001). *The Bootstrap in econometrics*, pp. 3159–3228. Elsevier Science B.V.
- Hsing, T. (1991). Extremal index estimation for a weekly dependent stationary sequence. *Annals of Statistics* 21(4), 1547–1569.
- Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics* 3, 456–499.
- Ibragimov, I. A. and Y. V. Linnik (1971). *Independent and stationary sequences of random variables*. The Netherlands: Walters-Noordhoff.
- Jaffe, J. F. and R. Westerfield (1985). The week-end effect in common stock returns: the international evidence. *Journal of Finance* 40, 433–454.
- Jansen, D. and C. de Vries (1991). On the frequency of large stock returns: Putting booms and busts into perspective. *Review of Economics and Statistics* 73, 19–24.
- Jeon, B. N. and T. C. Chiang (1991). A system of stock prices in world stock exchanges: Common stochastic trends for 1975-1990? *Journal of Economics and Business* 43, 329–338.
- Jones, G. L. (2004). On the markov chain central limit theorem. *Probability Surveys* 1, 299–320.
- Kaminski, G. and C. Reinhart (2000). On crises, contagion and confusion. *Journal of International Economics* 51(1), 145–168.
- Kaminski, G. L. and S. Schmukler (2008). Short-run pain, long-run gain: Financial liberalization and stock market cycles. *Review of Finance*. forthcoming.
- Kasa, K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics* 29, 95–124.
- Kim, S. J., F. Moshirian, and E. Wu (2005). Dynamic stock market integration driven by the european monetary union: An empirical analysis. *Journal of Banking and Finance* 29(10), 2475–2502.
- King, M. and S. Wadhwani (1990). Transmission of volatility between stock markets. *Review of Financial Studies* 3, 5–33.
- Klaassen, F. (2002). Improving garch volatility forecasts with regime-switching garch. *Empirical Economics* 27, 363–394.
- Koedijk, K. and M. Schaafgans (1973). Random difference equations and renewal theory for products of random matrices. *Acta Mathematica* 131, 207–248.



## BIBLIOGRAPHY

---

- Koedijk, K. and M. Schafgans (1990). The tail index of exchange rate returns. *Journal of International Economics* 29, 93–108.
- Koenker, R. and J. A. F. Machado (1999). Gmm inference when the number of moment conditions is large. *Journal of Econometrics* 93, 327–344.
- Korajczyk, R. A. (1996). Measure of stock market integration for developed and emerging markets. *World Bank Economic Review* 10(2), 267–289.
- Künsch, H. R. (1986). Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability* 23, 1025–1030.
- Lahaye, J., S. Laurent, and C. J. Neely (2007). Jumps, cojumps and macro announcements. Working paper.
- Lahiri, S. N. (1999). Theoretical comparisons of block bootstrap methods. *The Annals of Statistics* 27(1), 386–404.
- Lanne, M. (2007). Forecasting realized exchange rate volatility by decomposition. *International Journal of Forecasting* 23, 307–320.
- Ledford, A. and J. Tawn (1996). Statistics for near independence in multivariate extreme values. *Biometrika* 83(1), 169–187.
- Lien, D. and Y. K. Tse (2002). Some recent developments in futures hedging. *Journal of Economic Surveys* 16(3), 357–396.
- Lin, W., R. F. Engle, and T. Ito (1994). Do bulls and bears move across borders? international transmission of stock returns and volatility. *Review of Financial Studies* 7(3), 507–538.
- Lo, A. and J. Wang (2000). Trading volume: Definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies* 13, 257–300.
- Longin, F. and B. Solnik (2001). Extreme correlation of international equity markets. *Journal of Finance* 56, 649–676.
- Lowenstein, R. (2000). *When genius failed: The rise and fall of Long-Term Capital Management*. Random House.
- Lunde, A. and A. G. Timmermann (2004). Duration dependence in stock prices: An analysis of bull and bear markets. *Journal of Business and Economic Statistics* 22(3), 253–273.
- MacKenzie, D. (2003). Long-term capital management and the sociology of arbitrage. *Economy and Society* 32(3), 349–380.

- Maheu, J. M. and T. H. McCurdy (2000). Identifying bull and bear markets in stock returns. *Journal of Business and Economic Statistics* 18, 100–112.
- Malkiel, B. G. (1987). Efficient market hypothesis. In M. M. Eatwell, J. and P. Newman (Eds.), *The New Palgrave: A Dictionary of Economics*, pp. 120–123. London and New York: Macmillan and Stockton.
- Mallaris, A. G. and J. L. Urrutia (1992). The international crash of october 1987: Causality tests. *Journal of Financial and Quantitative Analysis* 27, 353–363.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7, 77–91.
- Martens, M. and D. van Dijk (2007). Measuring volatility with the realized range. *Journal of Econometrics* 138, 181–207.
- Masson, P. (1999). Contagion: Macroeconomic models with multiple equilibria. *Journal of International Money and Finance* 18, 587–602.
- Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics* 51, 247–257.
- Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. *Mitteilungsblatt für Mathematische Statistik* 8, 234–235.
- Newey, W. K. and K. D. West (1987). A simple, positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Obstfeld, M. (1996). Models of currency crises with self-fulfilling features. *European Economic Review* 40, 1037–1047.
- Obstfeld, M. and A. M. Taylor (2002). Globalization and capital markets. Working Paper 8846, National Bureau of Economic Research.
- Pagan, A. P. and K. A. Sossounov (2003). A simple framework for analysing bull and bear markets. *Journal of Applied Econometrics* 18, 233–246.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53, 61–65.
- Pericoli, M. and M. Sbracia (2003). A primer on financial contagion. *Journal of Economic Surveys* 17(4), 571–608.
- Piplack, J. and S. Straetmans (2008). Comovements of different asset classes during market stress. Working paper, Maastricht University.

## BIBLIOGRAPHY

---

- Poon, S.-H., M. Rockinger, and J. Tawn (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Review of Financial Studies* 17(2), 581–610.
- Quandt, R. (1960). Test of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55, 324–330.
- Quintos, C., Z. Fan, and P. Phillips (2001). Structural change tests in tail behaviour and the asian crisis. *Review of Financial Studies* 68, 633–663.
- Resnick, S. and C. Stărică (1998). Tail index estimation for dependent data. *The Annals of Applied Probability* 8(4), 1156–1183.
- Rogalski, R. (1984). New findings regarding day of the week returns over trading and non-trading periods. *Journal of Finance* 39, 1603–1614.
- Rogers, L. C. G. and S. E. Satchell (1991). Estimating variance from high, low, and closing prices. *Annals of Applied Probability* 1, 504–512.
- Roll, R. (1988). The international crash of october 1987. *Financial Analyst Journal* 44(5), 19–35.
- Rosenblatt, M. (1971). *Markov processes, structure and asymptotic behavior*. New York: Springer-Verlag.
- Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13, 341–360.
- Sabbatini, M. and O. Linton (1998). A garch model of the implied volatility of the swiss market index from option prices. *International Journal of Forecasting* 14, 199–213.
- Sachs, J., A. Tornell, and A. Velasco (1996). Financial crisis in emerging markets: the lessons from 1995. *Brookings Papers on Economic Activity* 1, 147–198.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* 6(2), 461–464.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19(3), 425–442.
- Shiller, R. (1989). Investor behavior in the october 1987 stock market crash: Survey evidence. In *Market Volatility*. Boston: Massachusetts Institute of Technology.
- Shiller, R. J. (2000). *Irrational exuberance*. Princeton: Princeton University Press.

- Siegel, J. J. (1998). *Stock for the long run*. New York: McGraw-Hill.
- Stentoft, L. (2008). Option pricing using realized volatility. Working paper, HEC Montréal, CREATES, CREF and CIRANO.
- Straetmans, S., W. Verschoor, and C. Wolff (2006). Extreme us stock market fluctuations in the wake of 9/11. *Journal of Applied Econometrics*. forthcoming.
- Summers, L. H. (1986). Does the stock market rationally reflect fundamental values? *Journal of Finance* 41, 591–602.
- Susmel, R. and R. F. Engle (1994). Hourly volatility spillovers between international equity markets. *Journal of International Money and Finance* 13(1), 3–25.
- Taylor, S. J. (1982). Financial returns modeled by the product of two stochastic processes - a study of daily sugar prices 1961-79. In O. D. Anderson (Ed.), *Time Series Analysis: Theory and Practice*, pp. 203–226. Amsterdam: North Holland.
- Teverovsky, V. and M. S. Taqqu (1997). Testing for long-range dependence in the presence of shifting means or a slowly-declining trend, using a varaince-type estimator. *Journal of Time Series Analysis* 18, 279–304.
- Yang, D. and Q. Zhang (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business* 73, 477–491.



---

## NEDERLANDSE SAMENVATTING (SUMMARY IN DUTCH)

### Het belang van kapitaalmarkten

Mondiale kapitaalmarkten zijn cruciaal voor de economie en hebben belangrijke functies. Nationale en internationale handel in beleggingsobjecten levert noodzakelijke ingrediënten, zoals (inter)nationale risicospreiding en intertemporele handel, voor efficiënte economische systemen. Daarom hebben niet alleen investeerders, maar ook politici en hele economieën belang bij een goede werking van kapitaalmarkten.

De aanhoudende globalisering heeft bijgedragen aan sterkere sociale, economische en financiële banden tussen landen door toenemende internationale handel en buitenlandse directe investeringen, de formatie van handelsblokken en geïntegreerde economische systemen zoals de EU, de opkomst van multinationale ondernemingen en financiële tussenpersonen, technologische en financiële innovaties, deregulatie en liberaliseringsprocessen.<sup>18</sup>

Zulke sterkere echte en financiële banden tussen economieën leiden meestal ook tot duidelijkere samenhang van hun kapitaalmarkten. Tegenwoordig kan in principe iedereen, binnen seconden, investeren in aandelen, obligaties en allerlei soorten beleggingsobjecten op iedere kapitaalmarkt ter wereld. Hierdoor kan nieuws in het ene deel van de wereld tegelijkertijd de prijsschommelingen van beleggingsobjecten in vele andere gebieden beïnvloeden. Over het algemeen zijn kapitaalmarkten dus van essentieel belang voor economieën. Deze markten komen bovendien in de hele wereld steeds meer met elkaar in verband te staan, waardoor externe factoren een steeds grotere invloed hebben op internationale economieën.<sup>19</sup>

---

<sup>18</sup>Zie, bijvoorbeeld, Jeon en Chiang (1991), en Chen, Firth, en Rui (2002).

<sup>19</sup>Over de impact van globalisering en economische integratie op kapitaalmarkten, zie, onder andere, Calvo en Reinhard (1996), Claessens, Dornbusch, en Park (2001), Forbes en Rigobon (2002), Obstfeld en Taylor (2002), en Goetzmann, Li, en Rouwenhorst (2005).

De moderne portefeuilletheorie, geïntroduceerd door Markowitz (1952), stelt voor hoe rationele investeerders diversificatie met imperfect gecorreleerde beleggingsobjecten zouden moeten gebruiken om hun portefeuille wat betreft zijn verwachte rendement en zijn risico of volatiliteit te optimaliseren.<sup>20</sup> Investeerders kunnen hun portefeuille internationaal en/of op het type beleggingsobjecten diversifiëren als dit een optimale risico-rendementverhouding schept. Om een optimale keuze voor beleggingsobjecten te maken is het essentieel om het parallelle verloop van internationale kapitaalmarkten, zoals rendementscorrelaties en volatiliteitsbewegingen, te bestuderen.

Behalve voor een optimale samenstelling van de portefeuille is een systematische analyse van de samenhang van kapitaalmarkten ook belangrijk vanwege de waarschijnlijke effecten van (extreme) prijsschommelingen van beleggingsobjecten op reële economische activiteit. Beleggingsobjecten zijn een van de belangrijkste factoren voor het kapitaal van mensen en daardoor voor het welzijn van hele economieën. Volgens de standaard economische theorie is vermogen een van de belangrijkste determinanten voor consumptie en investeringsbeslissingen. Ook veranderingen in de volatiliteit van beleggingsobjectprijzen kunnen een grote impact hebben op echte economische activiteit. Een verhoogde volatiliteit maakt het voor spelers moeilijker om hun toekomstige vermogensniveau te voorspellen en zou keuzes met betrekking tot consumptie, spaargeld en investeringen gemakkelijker kunnen vertekenen. Echte economische activiteit kan dus in hoge mate worden beïnvloed door schommelingen in de prijzen van beleggingsobjecten. Een beter begrip van zulke prijsveranderingen en volatiliteiten zou kunnen helpen om dergelijke effecten op te vangen.

Bovendien kunnen (negatieve) prijsveranderingen van beleggingsobjecten belangrijke effecten hebben op de stabiliteit van financiële systemen van vooral ontwikkelingslanden, maar ook van ontwikkelde landen. Een sterke negatieve schok op een kapitaalmarkt zou tot een verslechtering van de beleggingspositie kunnen leiden van, bijvoorbeeld, hedge funds. Dit zou mogelijk domino-effecten teweeg kunnen brengen bij door vreemdkapitaal gefinancierde banken en financiële systemen van hele economieën. Voorbeelden hiervan zijn de crisis van het Europees Monetair Systeem in 1992-93 en de ineenstorting van het *Long-Term Capital Management* (LTCM) hedge fund in 1998 (dat uiteindelijk in 2000 failliet ging). Beide gebeurtenissen waren de aanleiding voor onzekerheid en wereldwijd sterk volatiele kapitaalmarkten. De LTCM-crisis was ontstaan door grote investeringen gefinancierd met vreemdkapitaal en door ongunstige bewegingen op de kapitaalmarkt. Alleen nationale en internationale hulpprogramma's door nationale overheidsorganisaties als de Amerikaanse Centrale Bank (FED) en private banken konden

---

<sup>20</sup>Voor zijn bijdragen ontving Markowitz de Nobelprijs voor Economie in 1990.

voorkomen dat het fonds in 1998 instortte en tot mogelijke domino-effecten binnen de financiële sector zou leiden met grote economische consequenties.<sup>21</sup> Een recenter voorbeeld van gebeurtenissen met een wereldwijde economische impact zijn de subprime hypotheekcrises in 2007 en 2008.

Als kapitaalmarkten geen optimale resultaten boeken voor nationale economieën, zouden politici hun beleid met betrekking tot deze markten moeten aanpassen. Een prominent voorbeeld van beleid waarmee vele nationale overheden hebben getracht hun nationale kapitaalmarkten af te schermen voor internationale invloeden is controle op het kapitaalverkeer. Wetenschappers hebben nog geen consensus bereikt over de vraag of controle op het kapitaalverkeer de welvaart bevordert of niet.<sup>22</sup> Desondanks hangen conclusies over het nut van bepaalde maatregelen af van gedetailleerde kennis van de werking, dynamiek en samenhang van internationale financiële markten. Bij voorbeeld, negatief nieuws in het ene land of gebied kan ervoor zorgen dat investeerders hun beleggingen terugtrekken uit aanvankelijk onaangetaste markten, waardoor die landen ook met negatieve repercussies te maken krijgen. Politici moeten er dan naar streven om hun nationale kapitaalmarkt van internationale kapitaalmarkten te isoleren.

### **Doel van het onderzoek**

In de laatste decennia heeft onderzoek naar kapitaalmarkten en hun dynamiek en onderlinge verwevenheid aanzienlijke vooruitgang geboekt. Desondanks is er op dit gebied nog steeds ruimte voor vooruitgang omdat belangrijke kwesties zoals de cyclicaliteit van kapitaalmarkten nog niet in voldoende mate zijn behandeld. Daarnaast vragen financiële en overheidsinstellingen, maar ook investeerders, steeds vaker om geavanceerde instrumenten voor het modelleren van financiële markten en hun parallelle verloop, om hun beleid, investeringen of prijsberekeningen daarop te baseren.

Het parallelle verloop van kapitaalmarkten kan zich op verschillende wijzen manifesteren. Deze dissertatie verschaft een systematische benadering van verscheidene mogelijke dimensies van dit parallelle verloop. We werpen niet alleen een blik op de samenhang van beleggingsrendementen, maar presenteren ook een aantal andere manieren waarop dit fenomeen kan worden benaderd. Onder deze benaderingen vallen gelijktijdige sprongen van de prijzen van beleggingsobjecten, de cyclische samenloop van hausse- en baissemarkten, co-extreme rendementen of afhankelijkheid in de staarten tegenover rendementcorrelaties berekend over de gehele steekproef, en zowel het parallelle verloop van de volatiliteit van beleggingsrendementen als de samenloop van volatiliteitregimes van markten. Ook be-

---

<sup>21</sup>Zie Lowenstein (2000), Dunbar (2001), MacKenzie (2003), en Fenton-O'Creevy et al. (2004).

<sup>22</sup>Zie, bijvoorbeeld, De Gregorio et al. (2000).



handelen we de regionale dimensie (vooral Oost-Azië) versus mondiale samenloop. Over het geheel gezien is er duidelijk bewijs dat het parallelle verloop van prijsprocessen en -volatiliteiten van internationale kapitaalmarkten toeneemt, vooral in de laatste tien tot vijftien jaar.

Het doel van deze dissertatie is tweeledig. Ten eerste willen wij bijdragen aan de wetenschappelijke discussie over hoe bepaalde eigenschappen van internationale kapitaalmarkten met elkaar verbonden zijn en hoe deze verbindingen en parallel verloop zich door de jaren heen hebben ontwikkeld. We stellen onder andere een innovatieve benadering voor van het probleem van veranderende correlaties tussen kapitaalmarkten. Tot nu toe concentreerde wetenschappelijk onderzoek zich op pure rendementen met dagelijkse, wekelijkse of maandelijkse datafrequenties, waardoor de meting van pure correlatie gewoonlijk niet valide is vanwege de aanwezige heteroscedasticiteit. Wij kiezen er daarom voor om, in plaats van de pure rendementen zelf, eerst de cyclische component van rendementen af te leiden en deze vervolgens te analyseren. We tonen aan dat onderzoek naar veranderingen in het parallelle verloop van deze componenten in de ruwe data niet wordt vertekend door heteroscedasticiteit. Deze innovatieve aanpak om de cyclicaliteit van kapitaalmarkten te bestuderen voegt een nieuwe dimensie toe aan de bestaande kennis over parallel verloop van kapitaalmarkten en kan van belang zijn voor investeerders met een cyclische strategie of een investeringshorizon van middellange termijn.

Het tweede doel van deze dissertatie is om nieuwe econometrische instrumenten te introduceren die helpen ons begrip van het functioneren van kapitaalmarkten te vergroten en die ook onze capaciteiten verbeteren om de volatiliteit van kapitaalmarkten te voorspellen. We stellen een nieuw tijdreeksmodel voor om volatiliteit te kunnen schatten en voorspellen. Dit wordt vervolgens toegepast op data van aandelenmarkten om op het parallelle verloop van de volatiliteit van markten te controleren. We laten zien dat dit model beter presteert dan standaard GARCH-modellen. Daarnaast analyseren we het parallelle verloop van de gerealiseerde volatiliteit en co-jumps van aandelenmarkten van enkele geïndustrialiseerde landen.

Uit dit werk komt geen definitief antwoord voort op de vraag wat precies de oorzaak is van de gevonden toename van parallel verloop van kapitaalmarkten. Of zij veroorzaakt wordt door economische integratie of alleen door besmettelijke crises blijft een open vraag. In feite zijn er bewijzen voor beide verklaringen, afhankelijk van de markten en steekproeven waar we naar kijken. Desondanks zijn tijdreeksen tegenwoordig wellicht nog te kort om tot een duidelijk oordeel in het voordeel van een van beide hypothesen te komen. Daarom menen wij dat toekomstig onderzoek zich zou moeten richten op methodes die het mogelijk maken om een

onderscheid te maken tussen financiële integratie en besmetting als bronnen voor veranderende samenloop van kapitaalmarkten. Een goed uitgangspunt hiervoor is het werk van Candelon en Bodart (2004), die een frequentiedomeinbenadering voorstellen om de meer besmettelijke kortetermijneffecten en de langetermijneffecten van financiële integratie uit elkaar te halen. Zij beschouwen echter alleen de mogelijkheid van causale effecten tussen integratie en besmetting. Eén causaal verband tussen onderlinge afhankelijkheid en besmetting spreekt voor zich, aangezien sterkere financiële en economische verbindingen tot een grotere kwetsbaarheid voor asymmetrische shocks leiden. Het kan echter ook zo zijn dat besmetting tot een grotere integratie leidt gezien de gebruikelijke politieke reacties van landen bij het oplossen van een crisis.

### **Opbouw van de dissertatie**

In deze dissertatie komt in principe elk hoofdstuk, behalve de introductie in Hoofdstuk 1, overeen met een essay/paper. De originele papers zijn bijna identiek aan de hoofdstukken en dragen dezelfde titel. Aangezien de toepassingen, de gebruikte econometrische methodologieën en de focus per hoofdstuk/paper enigszins verschillen, worden deze in elk afzonderlijk hoofdstuk behandeld. De dissertatie bestaat in totaal uit zes hoofdstukken.

Hoofdstuk 2 analyseert gemeenschappelijke factoren in de continue component van volatiliteit, co-extreem- en co-jumpgedrag van een grotendeels Europese steekproef van aandelenmarktindices. Om die componenten in aandeelprijsprocessen op een handelsdag te identificeren, gebruiken we hoge-frequentiedata en -technieken. We laten zien dat één gemeenschappelijke factor in de meeste gevallen genoeg is om het grootste deel van de internationale variatie in het continue deel van volatiliteit te beschrijven en dat het belang van deze factor door de tijd heen is toegenomen. Bovendien vinden we sterke bewijzen voor asymmetrieën tussen extreem negatieve en positieve co-extreme rendementen uit sluitingsopeningskoersen en voor negatieve en positieve co-jumps tussen landen. Daarmee leidt Hoofdstuk 2 ook al enkele onderwerpen in die in de overige hoofdstukken van de dissertatie nog uitgebreider aan bod komen.

Hoofdstuk 3 implementeert schattings- en testprocedures voor parallel verloop van ‘cycli’ of ‘fases’ van aandelenmarkten in Azië. Hiervoor hebben we de *Generalized Method of Moments* (GMM)-benadering van Harding and Pagan (2006b) om de synchronisatie van conjunctuurcycli te meten uitgebreid, zodat we de cyclische synchronisatie van (bivariabele, multivariabele) aandelenmarkten kunnen schatten en testen in een kleine steekproef. We laten zien dat de asymptotische versie van de test in kleine steekproeven een steeds foutiever beeld geeft naarmate het aantal landen in de steekproef toeneemt. Een *block bootstrap*-versie van de test kan deze

vertekening echter verhelpen wanneer de lengte van de tijdreeks gedeeld door het aantal landen  $T/n$  groot genoeg is. Het kader van synchronisatie van financiële cycli hebben wij toegepast op vijf Oost-Aziatische landen. De nul-hypothese van een non-zero gemeenschappelijke synchronisatie-index kon niet verworpen worden voor bepaalde economisch belangrijke subsets van deze landen.

Hoofdstuk 4 bouwt voort op de ideeën uit Hoofdstuk 3 en poneert een test om structurele verandering in de bivariabele en multivariabele synchronisatie-indices endogeen vast te stellen. Uit de toepassing van deze test op vijf Aziatische aandelenmarkten komt een significante toename naar voren van parallel verloop tussen landen van Aziatische hausse- en baisseperiodes in 1997. Een studie naar de sterkte van de stabiliteitstest duidt erop dat de geconstateerde toename in parallel verloop meer van plotselinge aard is (i.e. besmetting of '*Asian Flu*') dan van graduele aard (i.e. financiële integratie). Daarnaast wordt gesteld dat de cycli van aandelenmarkten en hun neiging tot (verhoogde) synchronisatie nuttige informatie bevatten voor zowel investeerders, beleidsmakers als financiële regulerende instanties.

Hoofdstuk 5 berekent de samenhang tussen de belangrijkste typen financiële beleggingsobjecten van de Verenigde Staten (aandelen, obligaties, T-bills en goud) in periodes van financiële onrust. Onze resultaten hebben mogelijk belangrijke implicaties voor de strategische plaatsing van beleggingsobjecten en voor pensioenfondsenbeheer. We gebruiken multivariabele extreme waardetheorie om te schatten in welke mate een type beleggingsobject wordt blootgesteld aan extreme bewegingen in de andere typen objecten. Door structurele breuktesten toe te passen op die metingen onderzoeken we in hoeverre de samenhang tussen extreme beleggingsrendementen en volatiliteiten in verloop van tijd veranderen. Univariabele resultaten en bivariabele resultaten van parallel verloop vertonen significante breuken in de jaren '70 en '80 die overeenkomen met met turbulente periodes, zoals de oliecrises, Volcker's presidentschap van de FED of de beurskrach van 1987.

Hoofdstuk 6 brengt een nieuw model naar voren om volatiliteiten van kapitaalmarkten te beschrijven. Het stelt voor om de bandbreedte, dat wordt gedefinieerd als het verschil in percentage tussen de maximum- en minimumprijs van een beleggingsobject, te gebruiken als een niet-parametrische benadering van de volatiliteit van een beleggingsobject. We stellen vast dat tijdreeksen van bandbreedte heel goed kunnen worden gemodelleerd met als kader *Markov-switching* (MS) *autoregressive moving-average* (ARMA) met *generalized autoregressive conditional heteroscedasticity* (GARCH). Een dergelijk model maakt het mogelijk om verschillende volatiliteitsregimes te identificeren, i.e. een hoog en een laag volatiliteitsregime, en om periodes van volatiliteit  $h$  in de toekomst te voorspellen. In een internationale vergelijking vinden we bewijs voor parallel verloop van volatiliteit tussen verschillende landen.





---

## SHORT CURRICULUM VITAE

Jan Piplack was born on December 15, 1976 in Hamburg, Germany. He attended primary school from 1983 to 1987 and high school from 1987 to 1996 in Bargteheide, Germany. From high school he graduated in 1996 obtaining the Abitur (A-levels). He then proceeded for one year with the military service in the German Navy on Submarine U-23 stationed in Eckerförde, Germany. In 1997 he joined Deutsche Bank AG in Hamburg as an apprentice. After obtaining the degree Bankkaufmann (banker) with distinction in 1999 he started his study of International Economic Studies at Maastricht University, the Netherlands. During his study he worked as a teaching assistant giving classes in Quantitative Economics and various Bachelor and Master courses. He also studied two semesters at the University of Seville. In 2004 he obtained his Master's degree with distinction writing a thesis about volatility contagion.

After graduation Jan joined the Department of Economics at Maastricht University as a Ph.D. candidate under the supervision of Professors Bertrand Candelon and Joan Muysken until 2008. He also worked as a visiting researcher at the University of California in Santa Cruz, USA, from January to July 2007.

Jan has presented his work at numerous international conferences, among them the conference *Frontiers in Time Series Analysis*, and has been invited to present his work at various academic and professional institutions. His work has been published or is forthcoming in international refereed academic journals.